

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:30:15 ; Search time 21.5855 Seconds
(without alignments)
2286.817 Million cell updates/sec

Title: US-09-813-718-10_COPY_48_471

Perfect score: 2246
Sequence: 1 MSYKAAGBDYKADCPGNP.....VTDEIKEMTPRKLSDFRQ 424

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppa/US04_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubppa/US03_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppa/US02_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubppa/US01_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppa/US00_NEW_PUB.pep.*
9: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
11: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2246	100.0	437	9	US-09-813-718-12 Sequence 12, Appl
2	2246	100.0	471	9	US-10-126-4678-2 Sequence 2, Appl
3	2246	100.0	484	9	US-09-813-718-10 Sequence 10, Appl
4	2231	99.3	471	9	US-09-919-039-163 Sequence 163, Appl
5	2231	99.3	475	10	US-09-925-302-558 Sequence 558, Appl
6	2116	94.2	415	9	US-09-813-718-14 Sequence 14, Appl
7	1988	88.5	392	9	US-09-813-718-16 Sequence 16, Appl
8	1129.5	50.3	433	9	US-10-128-714-8545 Sequence 8545, Ap
9	831	37.0	173	10	US-09-925-302-855 Sequence 855, App
10	458	20.4	85	9	US-09-813-718-45 Sequence 45, Appl
11	419	18.7	179	9	US-10-128-714-3545 Sequence 3545, Ap
12	398	17.7	85	9	US-09-813-718-46 Sequence 46, Appl
13	363	16.2	85	9	US-09-813-718-48 Sequence 48, Appl
14	357	15.9	85	9	US-09-813-718-47 Sequence 47, Appl
15	356.5	15.9	142	10	US-09-925-302-557 Sequence 557, App
16	256	11.4	46	9	US-09-813-718-51 Sequence 51, Appl
17	210	9.3	46	9	US-09-813-718-52 Sequence 52, Appl
18	185.5	8.3	341	10	US-09-815-242-13444 Sequence 13444, A
19	180	8.0	46	9	US-09-813-718-54 Sequence 54, Appl

20	177	7.9	46	9	US-09-813-718-53 Sequence 53, Appl
21	147	6.5	385	9	US-10-128-714-3379 Sequence 3379, Ap
22	146.5	6.5	391	9	US-10-128-714-8379 Sequence 8379, Ap
23	146	6.5	339	10	US-09-815-242-11422 Sequence 11422, A
24	142.5	6.3	337	9	US-10-156-761-10954 Sequence 10954, A
25	134	6.0	372	9	US-09-813-718-8 Sequence 8, Appl
26	132.5	5.9	372	9	US-09-813-718-4 Sequence 4, Appl
27	132.5	5.9	536	9	US-09-813-718-2 Sequence 2, Appl
28	128	5.7	334	10	US-09-815-242-11070 Sequence 11070, A
29	113.5	5.1	423	10	US-09-815-242-10678 Sequence 10678, A
30	113	5.0	415	10	US-09-815-242-5494 Sequence 5494, Ap
31	113	5.0	420	10	US-09-815-242-12649 Sequence 12649, A
32	112	5.0	345	9	US-09-738-626-4352 Sequence 4252, Ap
33	110.5	4.9	331	9	US-10-156-761-12259 Sequence 12259, A
34	108	4.8	306	9	US-10-126-927-50 Sequence 50, Appl
35	108	4.8	306	9	US-10-126-927-59 Sequence 59, Appl
36	108	4.8	306	9	US-10-126-927-64 Sequence 64, Appl
37	108	4.8	306	9	US-10-126-931A-50 Sequence 50, Appl
38	108	4.8	306	9	US-10-126-931A-59 Sequence 59, Appl
39	108	4.8	306	9	US-10-126-931A-64 Sequence 64, Appl
40	108	4.8	334	10	US-09-815-242-13806 Sequence 13806, A
41	107	4.8	306	9	US-10-126-927-51 Sequence 51, Appl
42	107	4.8	306	9	US-10-126-931A-51 Sequence 51, Appl
43	106.5	4.7	334	10	US-09-815-242-10371 Sequence 10371, A
44	106	4.7	306	9	US-10-126-927-45 Sequence 45, Appl
45	106	4.7	306	9	US-10-126-931A-45 Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-09-813-718-12
Sequence 12, Application US/09813718
Publication No. US2002018266A1
GENERAL INFORMATION:
APPLICANT: Schimmel, Paul
TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
FILE REFERENCE: 00-221
CURRENT APPLICATION NUMBER: US/09/813,718
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 58
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 12
LENGTH: 437
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: human mini
OTHER INFORMATION: TPRS in PRT08
US-09-813-718-12
Query Match 100.0% Score 2246; DB 9; Length 437;
Best Local Similarity 100.0%; Pred. No. 2.1e-203; Mismatches 0; Indels 0; Gaps 0;
Matches 424; Conservative 0;
1 MSYKAAGBDYKADCPGNPAPTSNHPGDATEABEDFYDPMWVQTSKAGIDYKLIYRF 60
1 MSYKAAGBDYKADCPGNPAPTSNHPGDATEABEDFYDPMWVQTSKAGIDYKLIYRF 60
61 GSSKIDKELINIERATGQRPHFRGIFESHRRMNQVLDAYENKKPFYLYTGSPSE 120
61 GSSKIDKELINIERATGQRPHFRGIFESHRRMNQVLDAYENKKPFYLYTGSPSE 120
61 GSSKIDKELINIERATGQRPHFRGIFESHRRMNQVLDAYENKKPFYLYTGSPSE 120
121 AAHVCHLIPFITKMLQVFNPLVLTOMTDEKYLMLKQTLTDOAYGDAVENAKOIIACGF 180
121 AAHVCHLIPFITKMLQVFNPLVLTOMTDEKYLMLKQTLTDOAYGDAVENAKOIIACGF 180
121 AAHVCHLIPFITKMLQVFNPLVLTOMTDEKYLMLKQTLTDOAYGDAVENAKOIIACGF 180
181 DINKTFISDLDYMGSSGFYKNVVKIQKHVTFNOKVIFGFTSDCIKISFPAIOAAP 240
181 DINKTFISDLDYMGSSGFYKNVVKIQKHVTFNOKVIFGFTSDCIKISFPAIOAAP 240

QY 241 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALHSTFPALOGAQ 300
|
|
|
Db 241 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALHSTFPALOGAQ 300
QY 301 TKMSASDPNSSIFLITDTAKQIKTKVNHAFSGGSDTIEHRQFGGNCDDVVSFMYLTFEL 360
|
|
|
Db 301 TKMSASDPNSSIFLITDTAKQIKTKVNHAFSGGSDTIEHRQFGGNCDDVVSFMYLTFEL 360
QY 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAHQARKKVTDEIVKEFWTPRKL 420
|
|
|
Db 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAHQARKKVTDEIVKEFWTPRKL 420
QY 421 FDFQ 424
|
|
|
Db 421 FDFQ 424

RESULT 2

US-10-126-467B-2
; Sequence 2, Application US/10126467B
; Publication No. US2003005979A1
; GENERAL INFORMATION:
; APPLICANT: Paley, Elena
; TITLE OF INVENTION: MINIMAL MODEL OF AND TEST FOR ALZHEIMER'S DISEASE
; FILE REFERENCE: PALL-111
; CURRENT APPLICATION NUMBER: US/10/126,467B
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/284,980
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 09/513,895
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 09/384,869
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-126-467B-2

Query Match 100.0%; Score 2246; DB 9; Length 471;
Best Local Similarity 100.0%; Pred. No. 2.3e-203; Indels 0; Gaps 0;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSYKAAAGEDYKADCPGNPAPTSNHPDTEAEEDFVDPWTVQTSAGIDYDKLIVRF 60
|
|
|
Db 48 MSYKAAAGEDYKADCPGNPAPTSNHPDTEAEEDFVDPWTVQTSAGIDYDKLIVRF 107
QY 61 GSSKIDKELINRERATGQRPNHFLRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSSSE 120
|
|
|
Db 108 GSSKIDKELINRERATGQRPNHFLRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSSSE 167
QY 121 AMVGHILIPFIETKMLQDVFNPLVIOQMTDEKYLTKDLTDQAYGAVENAKDIACGF 180
|
|
|
Db 168 AMVGHILIPFIETKMLQDVFNPLVIOQMTDEKYLTKDLTDQAYGAVENAKDIACGF 227
QY 181 DINKTIFSDLDYMGSSGFYKVVVKIQKHVTFNQVGIFFGTFSDICIGISFPAIOAAP 240
|
|
|
Db 228 DINKTIFSDLDYMGSSGFYKVVVKIQKHVTFNQVGIFFGTFSDICIGISFPAIOAAP 287
QY 241 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALHSTFPALOGAQ 300
|
|
|
Db 241 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALHSTFPALOGAQ 300
QY 288 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALHSTFPALOGAQ 347
|
|
|
Db 301 TKMSASDPNSSIFLITDTAKQIKTKVNHAFSGGSDTIEHRQFGGNCDDVVSFMYLTFEL 360
|
|
|
Db 348 TKMSASDPNSSIFLITDTAKQIKTKVNHAFSGGSDTIEHRQFGGNCDDVVSFMYLTFEL 407
QY 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAHQARKKVTDEIVKEFWTPRKL 420
|
|
|
Db 408 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAHQARKKVTDEIVKEFWTPRKL 467

QY 421 FDFQ 424
|
|
|
Db 468 FDFQ 471

RESULT 3

US-09-813-718-10
; Sequence 10, Application US/09813718
; Publication No. US2002018266A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakasugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human
US-09-813-718-10

Query Match 100.0%; Score 2246; DB 9; Length 484;
Best Local Similarity 100.0%; Pred. No. 2.4e-203; Indels 0; Gaps 0;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSYKAAAGEDYKADCPGNPAPTSNHPDTEAEEDFVDPWTVQTSAGIDYDKLIVRF 60
|
|
|
Db 48 MSYKAAAGEDYKADCPGNPAPTSNHPDTEAEEDFVDPWTVQTSAGIDYDKLIVRF 107
QY 61 GSSKIDKELINRERATGQRPNHFLRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSSSE 120
|
|
|
Db 108 GSSKIDKELINRERATGQRPNHFLRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSSSE 167
QY 121 AMVGHILIPFIETKMLQDVFNPLVIOQMTDEKYLTKDLTDQAYGAVENAKDIACGF 180
|
|
|
Db 168 AMVGHILIPFIETKMLQDVFNPLVIOQMTDEKYLTKDLTDQAYGAVENAKDIACGF 227
QY 181 DINKTIFSDLDYMGSSGFYKVVVKIQKHVTFNQVGIFFGTFSDICIGISFPAIOAAP 240
|
|
|
Db 228 DINKTIFSDLDYMGSSGFYKVVVKIQKHVTFNQVGIFFGTFSDICIGISFPAIOAAP 287
QY 241 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALHSTFPALOGAQ 300
|
|
|
Db 241 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALHSTFPALOGAQ 300
QY 288 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALHSTFPALOGAQ 347
|
|
|
Db 301 TKMSASDPNSSIFLITDTAKQIKTKVNHAFSGGSDTIEHRQFGGNCDDVVSFMYLTFEL 360
|
|
|
Db 348 TKMSASDPNSSIFLITDTAKQIKTKVNHAFSGGSDTIEHRQFGGNCDDVVSFMYLTFEL 407
QY 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAHQARKKVTDEIVKEFWTPRKL 420
|
|
|
Db 408 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAHQARKKVTDEIVKEFWTPRKL 467
QY 421 FDFQ 424
|
|
|
Db 468 FDFQ 471

RESULT 4

US-09-919-039-163
; Sequence 163, Application US/09919039
; Publication No. US2003010887A1
; GENERAL INFORMATION:
; APPLICANT: Kaefer, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES

FILE REFERENCE: PA-0035 US
CURRENT APPLICATION NUMBER: US/09/919,039
CURRENT FILING DATE: 2002-09-09
PRIOR APPLICATION NUMBER: 60/222,113
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 401
SOFTWARE: PERL Program
SEQ ID NO 163
LENGTH: 471
TYPE: PRF
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030108871A1 2705515CD1
US-09-919-039-163

Query Match 99.3%; Score 2231; DB 9; Length 471;
Best Local Similarity 99.5%; Pred. No. 6.1e-202;
Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSYKAAAGEDYKADCPGPNPAPTSNHGPDATAEEDFVDPMTVQSSAKGIDYDKLIYRF 60
DB 48 MSYKAAAGEDYKADCPGPNPAPTSNHGPDATAEEDFVDPMTVQSSAKGIDYDKLIYRF 107
QY 61 GSSKIDKELINRERATGQRPHHFLRGIFFSHRDMNVLDAYENKKPFYLYTGRGSSSE 120
DB 108 GSSKIDKELINRERATGQRPHHFLRGIFFSHRDMNVLDAYENKKPFYLYTGRGSSSE 167
QY 121 AMHVGHLPFTTKMLQDVFNVPVLYQWTDDEKYLKMLTLDOAYGDAVENAKDIIACGF 180
DB 166 AMHVGHLPFTTKMLQDVFNVPVLYQWTDDEKYLKMLTLDOAYGDAVENAKDIIACGF 227
QY 181 DINKTFISDDLDMYMGSSGFYKNVVKIQKHVTFNQVKGIFGFTSDCIKISFPALQAP 240
DB 228 DINKTFISDDLDMYMGSSGFYKNVVKIQKHVTFNQVKGIFGFTSDCIKISFPALQAP 287
QY 241 SFSNSFPQIFRDRTDIQCILPCALIDDPYFRMTRDVAPRIGYKPKALLHSTFFPALQAG 300
DB 288 SFSNSFPQIFRDRTDIQCILPCALIDDPYFRMTRDVAPRIGYKPKALLHSTFFPALQAG 347
QY 301 TMSASDPNSSIFLTDYKQIKTKVKNKAFSGGRDTIEHRQFGNCVDVDSFMYLTFFL 360
DB 348 TMSASDPNSSIFLTDYKQIKTKVKNKAFSGGRDTIEHRQFGNCVDVDSFMYLTFFL 407
QY 361 EDDDKLEQIRKDYTGAMLTGELKKALIEVLOPLAEHQARRKVEYDEIVKGFMTPRKLS 420
DB 408 EDDDKLEQIRKDYTGAMLTGELKKALIEVLOPLAEHQARRKVEYDEIVKGFMTPRKLS 467
QY 421 PDPQ 424
DB 468 PDPQ 471

RESULT 5
US-09-925-302-558
Sequence 558, Application US/09925302
Patent No. US20020044941A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA104
CURRENT APPLICATION NUMBER: US/09/925,302
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05918
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
NUMBER OF SEQ ID NOS: 896
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 558
LENGTH: 475
TYPE: PRF
ORGANISM: Homo sapiens

US-09-925-302-558

Query Match 99.3%; Score 2231; DB 10; Length 475;
Best Local Similarity 99.5%; Pred. No. 6.2e-202;
Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSYKAAAGEDYKADCPGPNPAPTSNHGPDATAEEDFVDPMTVQSSAKGIDYDKLIYRF 60
DB 52 MSYKAAAGEDYKADCPGPNPAPTSNHGPDATAEEDFVDPMTVQSSAKGIDYDKLIYRF 111
QY 61 GSSKIDKELINRERATGQRPHHFLRGIFFSHRDMNVLDAYENKKPFYLYTGRGSSSE 120
DB 112 GSSKIDKELINRERATGQRPHHFLRGIFFSHRDMNVLDAYENKKPFYLYTGRGSSSE 171
QY 121 AMHVGHLPFTTKMLQDVFNVPVLYQWTDDEKYLKMLTLDOAYGDAVENAKDIIACGF 180
DB 172 AMHVGHLPFTTKMLQDVFNVPVLYQWTDDEKYLKMLTLDOAYGDAVENAKDIIACGF 231
QY 181 DINKTFISDDLDMYMGSSGFYKNVVKIQKHVTFNQVKGIFGFTSDCIKISFPALQAP 240
DB 232 DINKTFISDDLDMYMGSSGFYKNVVKIQKHVTFNQVKGIFGFTSDCIKISFPALQAP 291
QY 241 SFSNSFPQIFRDRTDIQCILPCALIDDPYFRMTRDVAPRIGYKPKALLHSTFFPALQAG 300
DB 292 SFSNSFPQIFRDRTDIQCILPCALIDDPYFRMTRDVAPRIGYKPKALLHSTFFPALQAG 351
QY 301 TMSASDPNSSIFLTDYKQIKTKVKNKAFSGGRDTIEHRQFGNCVDVDSFMYLTFFL 360
DB 352 TMSASDPNSSIFLTDYKQIKTKVKNKAFSGGRDTIEHRQFGNCVDVDSFMYLTFFL 411
QY 361 EDDDKLEQIRKDYTGAMLTGELKKALIEVLOPLAEHQARRKVEYDEIVKGFMTPRKLS 420
DB 412 EDDDKLEQIRKDYTGAMLTGELKKALIEVLOPLAEHQARRKVEYDEIVKGFMTPRKLS 471
QY 421 PDPQ 424
DB 472 PDPQ 475

RESULT 6
US-09-813-718-14
Sequence 14, Application US/09813718
Publication No. US2002018266A1
GENERAL INFORMATION:
APPLICANT: Schimmel, Paul
APPLICANT: Wakasugi, Keisuke
TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
FILE REFERENCE: 00-221
CURRENT APPLICATION NUMBER: US/09/813,718
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 415
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: human
US-09-813-718-14

Query Match 94.2%; Score 2116; DB 9; Length 415;
Best Local Similarity 100.0%; Pred. No. 3.6e-191;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 SNHGDATEAEDFVDPMTVQSSAKGIDYDKLIYRFSSKIDKELINRERATGQRPHH 83
DB 2 SNHGDATEAEDFVDPMTVQSSAKGIDYDKLIYRFSSKIDKELINRERATGQRPHH 61
QY 84 FLRGIFFSHRDMNVLDAYENKKPFYLYTGRGSSSEAMHVGHLPFTTKMLQDVFNVP 143
DB 62 FLRGIFFSHRDMNVLDAYENKKPFYLYTGRGSSSEAMHVGHLPFTTKMLQDVFNVP 121

```

QY 144 LV1QMTDEKTLAKDLTDQAYGDAVENAKDIIACGPDINKTIFPSDLDMGSSGGYKN 203
DB 122 LV1QMTDEKTLAKDLTDQAYGDAVENAKDIIACGPDINKTIFPSDLDMGSSGGYKN 181
QY 204 VK1QKAVTNOVGIFGFTSDSCIGKISPPAIAAASFSNFPQIFPRDRIDICLIPCA 263
DB 182 VK1QKAVTNOVGIFGFTSDSCIGKISPPAIAAASFSNFPQIFPRDRIDICLIPCA 241
QY 264 IDDPYFRMRDVAIRIGYPPKALLHSTFPFALQAGQTKMSASDPNSSIFLTDIAKQIKT 323
DB 242 IDDPYFRMRDVAIRIGYPPKALLHSTFPFALQAGQTKMSASDPNSSIFLTDIAKQIKT 301
QY 324 KVNHAASGGRTIEEHRQCGNDVVSFWYLFPLEDDDKLEQIRKDYTGSGMLTGL 383
DB 302 KVNHAASGGRTIEEHRQCGNDVVSFWYLFPLEDDDKLEQIRKDYTGSGMLTGL 361
QY 384 KKALIEVLQPLIAHQAARKEVTDEIYKEFWTPKLSFDFQ 424
DB 362 KKALIEVLQPLIAHQAARKEVTDEIYKEFWTPKLSFDFQ 402

```

RESULT 7

```

US-09-813-718-16
; Sequence 16, Application US/09813718
; Publication No. US20020182666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakasugi, Keisuke
; TITLE OF INVENTION: Human Antinociceptive RNA Synthesase Polypeptides Useful For
; TITLE OF INVENTION: The Regulation of Angiogenesis
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813, 718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 16
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human minor
; OTHER INFORMATION: Trps fragment in pr20B
US-09-813-718-16

```

Query Match 88.5%; Score 1988; DB 9; Length 392;
Best Local Similarity 100.0%; Pred. No. 3,9e-179;

Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 47 SAKGIDVDKLVIRGSSKIDKELNRLERATGQRPHHFLRGIFFSHRDMNOVLDAYENK 106
DB 2 SAKGIDVDKLVIRGSSKIDKELNRLERATGQRPHHFLRGIFFSHRDMNOVLDAYENK 61
QY 107 KPFLYLRGSSSEAMVHGLIPFTKMLQDVNVNVLV1QMTDEKTLAKDLTDQAYG 166
DB 62 KPFLYLRGSSSEAMVHGLIPFTKMLQDVNVNVLV1QMTDEKTLAKDLTDQAYG 121
QY 167 DAVENAKDIIACGPDINKTIFPSDLDMGSSGGYKVVVKIQKAVTNOVGIFGFTSD 226
DB 122 DAVENAKDIIACGPDINKTIFPSDLDMGSSGGYKVVVKIQKAVTNOVGIFGFTSD 181
QY 227 CIGKISPPAIAAASFSNFPQIFPRDRIDICLIPCAIDDPYFRMRDVAIRIGYPPK 286
DB 182 CIGKISPPAIAAASFSNFPQIFPRDRIDICLIPCAIDDPYFRMRDVAIRIGYPPK 241
QY 287 LHSSTFPALQAGQTKMSASDPNSSIFLTDIAKQIKTKVNHAFSSGGRDTIEEHRQCGN 346
DB 242 LHSSTFPALQAGQTKMSASDPNSSIFLTDIAKQIKTKVNHAFSSGGRDTIEEHRQCGN 301
QY 347 CDVVSFWYLFPLEDDDKLEQIRKDYTGSGMLTGLKKALIEVLQPLIAHQAARKEVT 406
DB 302 CDVVSFWYLFPLEDDDKLEQIRKDYTGSGMLTGLKKALIEVLQPLIAHQAARKEVT 361

```

```

QY 407 DEIVKEFWTPKLSFDFQ 424
DB 362 DEIVKEFWTPKLSFDFQ 379

```

RESULT 8

```

US-10-128-714-8545
; Sequence 8545, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengji
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroskinin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128, 714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285, 697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287, 066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295, 890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303, 899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316, 362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 8545
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8545

```

Query Match 50.3%; Score 1129.5; DB 9; Length 433;
Best Local Similarity 53.7%; Pred. No. 4.4e-98;
Matches 227; Conservative 64; Mismatches 115; Indels 17; Gaps 7;

```

QY 13 ADGPPNPAPT-----SNHGPDATAEEDFVDWMTV-----QTSAGKIDVDKLVIRGSS 63
DB 2 ADGPPNPAPTLTLISLEKAP-ASKAVAYVTFPVDVSGVDSGKLLPVYDKLVREFGAT 60
QY 64 KIDKELNRLERATGQRPHHFLRGIFFSHRDMNOVLDAYENKPPFLYLRGSSSEAMH 123
DB 61 RISKELERFERVTGRRPHRFRGIFFSHRDMNLILDRYKQGPFLYLRGSSSDSMH 120
QY 124 VGHILPITTKMLQDVNVNVLV1QMTDEKTLAKDLTDQAYGDAVENAKDIIACGPD 182
DB 121 VGHITPFEFTKMLQDVNVNVLV1QMTDEKTLAKDLTDQAYGDAVENAKDIIACGPD 180
QY 183 NKFEITPSDLDMGSSGGYKVVVKIQKAVTNOVGIFGFTSDSCIGKISPPAIAAASFS 242
DB 181 KTFITPSDLPVGG--CAFEENICRNKAKRITINSVKTFFPDSNNVGEHFFCATGATATF 238
QY 243 SNFSPQIF-RDR---TDIOCLIPCAIDDPYFRMRDVAIRIGYPPKALLHSTFPALQ 298
DB 239 AFSFPHIFGDRKVVSIICLIPCAIDDPYFRQCEHAKKKVKKPSLIHAFIPALQ 298
QY 299 AQTQMSASDPNSSIFLTDIAKQIKTKVNHAFSSGGRDTIEEHRQCGNDVVSFWYLF 358
DB 299 PGSKMSASVETSAIFMNDAPNRIKNKINKYAFSSGGDTAELOKATGANTKDVPPFOYLT 358
QY 359 FLEDDDKLEQIRKDYTGSGMLTGLKKALIEVLQPLIAHQAARKEVTDEIYKEFWTPK 418
DB 359 FWEDEELERIVAYEKGEMLTGEVAKQCIATLOAYVQAFERRAVTDEIYAEFWPNS 418
QY 419 LSF 421
DB 419 LSF 421

```


Db 419 LEW 421

RESULT 9

US-09-925-302-855
; Sequence 855, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 855
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (159)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; LOCATION: (168)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-855

Query Match 37.0%; Score 831; DB 10; Length 173;

Best Local Similarity 98.7%; Pred. No. 1.7e-70; Mismatches 2; Indels 0; Gaps 0;

Matches 156; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 130 FIFTKMLQDVFNPLVYQMTDDEKYLKMLTLPOAGDAVENAKDIIACGFDINKTFPS 189
DB 1 FIFTKMLQDVFNPLVYQMTDDEKYLKMLTLPOAGDAVENAKDIIACGFDINKTFPS 60
QY 190 DLDYMGSSGFFYNVVKIQHVTENOVKGIFFGFTSDSCIGKISPPAIOAAPSFSNSFPQI 249
DB 61 DLDYMGSSGFFYNVVKIQHVTENOVKGIFFGFTSDSCIGKISPPAIOAAPSFSNSFPQI 120
QY 250 FRDRTDICLICPAIDODPYFRMTDVPARIGYKPKAL 287
DB 121 FRDRTDICLICPAIDODPYFRMTDVPARIGYKPKAL 158

RESULT 10

US-09-813-718-45
; Sequence 45, Application US/09813718
; Publication No. US20020182666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Makasugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-813-718-45

Query Match 20.4%; Score 458; DB 9; Length 85;
Best Local Similarity 100.0%; Pred. No. 9.5e-36;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSYKAAGBDYKADCCPGNPAPTSNHPDTEAEEDFVDPMWTVQTSASAGIDYDKLIVRF 60

Db 1 MSYKAAGBDYKADCCPGNPAPTSNHPDTEAEEDFVDPMWTVQTSASAGIDYDKLIVRF 60

QY 61 GSSKIDKELINRIERATGQRPHHFL 85
DB 61 GSSKIDKELINRIERATGQRPHHFL 85

RESULT 11

US-10-128-714-3545
; Sequence 3545, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengli
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Broshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3545
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3545

Query Match 18.7%; Score 419; DB 9; Length 179;
Best Local Similarity 55.7%; Pred. No. 1.3e-31;
Matches 83; Conservative 21; Mismatches 35; Indels 10; Gaps 3;

QY 13 ADCPCNPAPV-----SNHPDTEAEEDFVDPMWTV-----OTSASAGIDYDKLIVRFSS 63
DB 2 ADCPCNPAPVLTLSLAERAP-ASKAVAQVTFPVDVSGVDSGKLLPVYDKLVREFGAT 60
QY 64 KIDKELINRIERATGQRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRSPSEAMH 123
DB 61 RISKELERFERVTGRRPFRMRGIVFSHRDLNLLDREYEGQPPYLYTGTGSPSSDSMH 120
QY 124 VGHLPFIFTKMLQDVFNPLVYQMTDDE 152
DB 121 VGHLPFIFTKMLQDVFNPLVYQMTDDE 149

RESULT 12

US-09-813-718-46
; Sequence 46, Application US/09813718
; Publication No. US20020182666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Makasugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58

SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 46
 LENGTH: 85
 TYPE: PRT
 ORGANISM: Bos taurus
 US-09-813-718-46

Query Match
 Best Local Similarity 88.1%; Score 398; DB 9; Length 85;
 Matches 74; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 SYKAAAGEDYKADCPGPNPAPTSNHPDTEAEEDFVDPMTVQTSAGIDYDKLIVRF 61
 DB 2 SYKAAAGEDYKADCPGPNPAPTSNHPDTEAEEDFVDPMTVQTSAGIDYDKLIVRF 61
 DB 62 GSKIDKELINRIERATGQRPVHFL 85
 DB 62 GSKIDKELINRIERATGQRPVHFL 85

RESULT 13

US-09-813-718-48
 Sequence 48; Application US/09813718
 Publication No. US2002018266A1
 GENERAL INFORMATION:
 APPLICANT: Schimmel, Paul
 APPLICANT: Makasugi, Keisuke
 TITLE OF INVENTION: Human Antinuclyl-cRNA Synthetase Polypeptides Useful For
 TITLE OF INVENTION: The Regulation of Angiogenesis
 FILE REFERENCE: 00-221
 CURRENT APPLICATION NUMBER: US/09/813, 718
 CURRENT FILING DATE: 2001-03-21
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 48
 LENGTH: 85
 TYPE: PRT
 ORGANISM: Oryctolagus cuniculus
 US-09-813-718-48

Query Match
 Best Local Similarity 82.4%; Score 363; DB 9; Length 85;
 Matches 70; Conservative 5; Mismatches 8; Indels 2; Gaps 2;

QY 2 SYKAAAGEDYKADCPGPNPAPTSNHPD-ATEAEEDFVDPMTVQTSAGIDYDKLIVRF 60
 DB 2 SYKAAAGEDYKADCPGPNPAPTSNHPD-ATEAEEDFVDPMTVQTSAGIDYDKLIVRF 60
 QY 61 GSKIDKELINRIERATGQRPVHFL 85
 DB 61 GSKIDKELINRIERATGQRPVHFL 85

RESULT 14
 US-09-813-718-47
 Sequence 47; Application US/09813718
 Publication No. US2002018266A1
 GENERAL INFORMATION:
 APPLICANT: Schimmel, Paul
 APPLICANT: Makasugi, Keisuke
 TITLE OF INVENTION: Human Antinuclyl-cRNA Synthetase Polypeptides Useful For
 TITLE OF INVENTION: The Regulation of Angiogenesis
 FILE REFERENCE: 00-221
 CURRENT APPLICATION NUMBER: US/09/813, 718
 CURRENT FILING DATE: 2001-03-21
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 47
 LENGTH: 85
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-813-718-47

Query Match
 Best Local Similarity 81.2%; Score 357; DB 9; Length 85;
 Matches 69; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSYKAAAGEDYKADCPGPNPAPTSNHPDTEAEEDFVDPMTVQTSAGIDYDKLIVRF 60
 DB 1 MSYKAAAGEDYKADCPGPNPAPTSNHPDTEAEEDFVDPMTVQTSAGIDYDKLIVRF 60
 QY 61 GSKIDKELINRIERATGQRPVHFL 85
 DB 61 GSKIDKELINRIERATGQRPVHFL 85

RESULT 15

US-09-925-302-557
 Sequence 557; Application US/09925302
 Patent No. US20020044941A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 FILE REFERENCE: PA104
 CURRENT APPLICATION NUMBER: US/09/925,302
 CURRENT FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US00/05918
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/124,270
 PRIOR FILING DATE: 1999-03-12
 NUMBER OF SEQ ID NOS: 896
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 557
 LENGTH: 142
 TYPE: PRT
 ORGANISM: Homo sapiens

FEATURE:
 NAME/KEY: SITE
 LOCATION: (115)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (117)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (122)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (124)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (130)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (137)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (142)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-925-302-557

Query Match
 Best Local Similarity 78.0%; Score 356.5; DB 10; Length 142;
 Matches 71; Conservative 5; Mismatches 12; Indels 3; Gaps 2;

QY 1 MSYKAAAGEDYKADCPGPNPAPTSNHPDTEAEEDFVDPMTVQTSAGIDYDKLIVRF 60
 DB 52 MSYKAAAGEDYKADCPGPNPAPTSNHPDTEAEEDFVDPMTVQTSAGIDYDKLIVRF 111
 QY 61 GSKIDKELINRIERATGQRPVHFLRGIIF 91
 DB 112 GSSXNKOELIXDE-STAXKTHS--GQGXFF 139

Search completed: July 10, 2003, 12:33:36
 Job time: 22.5855 secs

RESULT 1
AAB47616
ID AAB47616 standard; Protein; 437 AA.
XX
AC AAB47616;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human mini TtpRS.
XX
KW Tyrosyl-tRNA synthetase; TyrRS; Rosemann fold nucleotide binding domain
KW vascular endothelial cell function; burn; plastic surgery; abdomen;
KW polyporphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing
KW dermal ulcer; diabetic ulcer; endothelialization;
XX tryptophanyl-tRNA synthetase; ttpRS; vascular graft surgery.
OS Homo sapiens.
XX
FN WC200174841-A1.
XX
PD 11-OCT-2001.
XX
PF 21-MAR-2001; 2001WC-US08966.
XX
PR 31-MAR-2000; 2000US-193471P.
XX
PA (SCRI.) SCRIPPS RES INST.
XX
PI Schimmel P, Wakaugi K;
XX
DR WPJ; 2001-626377/72.
DR N-PsDB; AAA43603.


```

Db 61 GSSKIDKELINRERATGGRPHFLRGRIFFSHRDMNOVLDAVENKKPFYLTGRGPSS 120
QY 121 AMHVGHLIPFIPTKMLQDVFNPLVIQMTDDEKYLKMDLTLDQAGDAVENAKDIIACGF 180
Db 121 AMHVGHLIPFIPTKMLQDVFNPLVIQMTDDEKYLKMDLTLDQAGDAVENAKDIIACGF 180
QY 181 DINKTFIFSDLDYMGSSGFYKNNVKIQKHTFNQYKIGFTDSDICIGISFPALQAP 240
Db 181 DINKTFIFSDLDYMGSSGFYKNNVKIQKHTFNQYKIGFTDSDICIGISFPALQAP 240
QY 241 SFSNSFPQIFRDRDITQCLIPCAIDDPYFRMTRDVAIRIGYKPPALHSTFPALQAG 300
Db 241 SFSNSFPQIFRDRDITQCLIPCAIDDPYFRMTRDVAIRIGYKPPALHSTFPALQAG 300
QY 301 TKMSASDPNSSIFLDTAKQIKTKNNKHAFFSGGRDTEBHRQFGNCDDVVSFMYLTFFL 360
Db 301 TKMSASDPNSSIFLDTAKQIKTKNNKHAFFSGGRDTEBHRQFGNCDDVVSFMYLTFFL 360
QY 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLS 420
Db 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLS 420
QY 421 PDFQ 424
Db 421 PDFQ 424

```

RESULT 3

AA047615
ID AA047615 standard; Protein; 484 AA.

XX AA047615;

DT 07-JAN-2002 (first entry)

DE Human full-length TrpRS.

XX Tyrosyl-tRNA synthetase; TrpRS; Roessmann fold nucleotide binding domain;
KW vascular endothelial cell function; burn; plastic surgery; abdomen;
KW polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
KW dermal ulcer; diabetic ulcer; endothelialization;
KW tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.

OS Homo sapiens.

PN WO200174841-A1.

PD 11-OCT-2001.

PF 21-MAR-2001; 2001MO-US08966.

PR 31-MAR-2000; 2000US-193471P.

PA (SCRI) SCRIIPS RRS INST.

PI Schimmel P, Wakasugi K;

DR WPI; 2001-626377/72.

DR N-PSDB; AA043602.

XX New human truncated tyrosyl-tRNA synthetase polypeptide for regulating

XX PT vascular endothelial function, in particular for regulating

XX PT angiogenesis, tumor metastasis and treating myocardial infarction

XX PS Disclosure; Page 117-19; 150pp; English.

CC The sequences given in AA047615-18 show full length and truncated
CC versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
CC the invention comprises a Roessmann fold nucleotide binding domain, and
CC is capable of regulating vascular endothelial cell function. It is of
CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of

CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,
CC enhancing angiogenesis to a graft, treating myocardial infarction,
CC solid tumor, and a condition that would benefit from increased or
CC decreased angiogenesis in a mammal, in particular humans. It is also
CC useful in diagnosis and as a wound healing agent for treating wounds
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
CC plastic surgery when reconstruction is required following a burn or
CC for cosmetic purposes. It is particularly useful in the treatment of
CC abdominal wounds where there is high risk of infection. Truncated TrpRS
CC promotes endothelialization in vascular graft surgery and is used in
CC conjunction with angiography to administer the angiogenic RNA
CC synthetase polypeptides or polynucleotides directly to the lumen and
CC wall of the blood vessel.

SO Sequence 484 AA;

Query Match 100.0%; Score 2246; DB 22; Length 484;

Best Local Similarity 100.0%; Pred. No. 8.7e-222;

Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSYKAAAGDYADCEPQNPAPTSNHPDAAEDFVDPMVTQSSAKGIDYDKLYVF 60
Db 48 MSYKAAAGDYADCEPQNPAPTSNHPDAAEDFVDPMVTQSSAKGIDYDKLYVF 107
QY 61 GSSKIDKELINRERATGGRPHFLRGRIFFSHRDMNOVLDAVENKKPFYLTGRGPSS 120
Db 108 GSSKIDKELINRERATGGRPHFLRGRIFFSHRDMNOVLDAVENKKPFYLTGRGPSS 167
QY 121 AMHVGHLIPFIPTKMLQDVFNPLVIQMTDDEKYLKMDLTLDQAGDAVENAKDIIACGF 180
Db 168 AMHVGHLIPFIPTKMLQDVFNPLVIQMTDDEKYLKMDLTLDQAGDAVENAKDIIACGF 227
QY 181 DINKTFIFSDLDYMGSSGFYKNNVKIQKHTFNQYKIGFTDSDICIGISFPALQAP 240
Db 228 DINKTFIFSDLDYMGSSGFYKNNVKIQKHTFNQYKIGFTDSDICIGISFPALQAP 287
QY 241 SFSNSFPQIFRDRDITQCLIPCAIDDPYFRMTRDVAIRIGYKPPALHSTFPALQAG 300
Db 288 SFSNSFPQIFRDRDITQCLIPCAIDDPYFRMTRDVAIRIGYKPPALHSTFPALQAG 347
QY 301 TKMSASDPNSSIFLDTAKQIKTKNNKHAFFSGGRDTEBHRQFGNCDDVVSFMYLTFFL 360
Db 348 TKMSASDPNSSIFLDTAKQIKTKNNKHAFFSGGRDTEBHRQFGNCDDVVSFMYLTFFL 407
QY 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLS 420
Db 408 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLS 467
QY 421 PDFQ 424
Db 468 PDFQ 471

```

RESULT 4

AA013491
ID AA013491 standard; Protein; 484 AA.

XX AA013491;

DT 12-FEB-2002 (first entry)

DE Human tryptophanyl t-RNA synthetase (TrpRS) in pET20B.

XX Human; tryptophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;

XX KW TrpRS; vascular endothelial cell function; angiogenesis; wound healing;

XX KW re-vascularization; dermal ulcer; pressure sore; venous ulcer; injury;

XX KW diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;

XX KW angiotensin; gene therapy; tumour; inflammation; vascular permeability;

XX KW Rheumatoid arthritis; psoriasis; diabetic retinopathy.

XX OS Homo sapiens.

PN WO200175078-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 21-MAR-2001; 2001WO-US08975.
 XX
 PR 31-MAR-2000; 2000US-193471P.
 XX (SCRI) SCRI:PPS RES INST.
 PA
 PI Schimmel P, Wakasugi K;
 DR WPI; 2002-010784/01.
 XX N-PSDB; AAD22482.
 PT
 PT Novel truncated tryptophanyl-tRNA synthetase polypeptides capable of
 PT regulating vascular endothelial cell function, preferably angiogenesis,
 PT is useful for treating solid tumor or suppressing tumor metastasis in
 PT mammal -

PS Example 1; Page 117-119; 149pp; English.

XX The patent discloses human aminoacyl tRNA synthetases, particularly
 CC truncated tryptophanyl-tRNA synthetases (TrpRS) comprising a Rossmann
 CC fold nucleotide binding domain and polynucleotides encoding them. The
 CC invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS
 CC sequences are useful for regulating vascular endothelial cell function,
 CC preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound
 CC healing agents for re-vascularizing damaged tissues. They are useful for
 CC treating full-thickness wounds (e.g. dermal ulcers, including pressure
 CC sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS
 CC sequences can also be used in plastic surgery when reconstruction is
 CC required following a burn, other trauma, or even for cosmetic purposes.
 CC Angiogenic TrpRS is also used in association with surgery and following
 CC the repair of cuts, for promoting endothelialisation in vascular graft
 CC surgery and for repairing the damage of myocardial infarction and in
 CC conjunction with coronary bypass surgery by stimulating the growth of
 CC transplanted tissue. TrpRS is also used in conjunction with angiography.
 CC TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in
 CC immunoassays to detect the presence of tumours. They are also useful
 CC for blocking endogenous angiogenic activity and retard the growth of
 CC solid tumours. These antibodies may also be used to treat inflammation
 CC caused by increased vascular permeability. Inhibiting the activity of
 CC TrpRS by antisense technology is useful for preventing further growth
 CC or even regress solid tumours, and for treating rheumatoid arthritis,
 CC psoriasis, diabetic retinopathy, all of which are characterised by
 CC abnormal angiogenesis. The present sequence is human tryptophanyl
 CC t-RNA synthetase (TrpRS) in PET20B.

XX Sequence 484 AA;

Query Match 100.0%; Score 2246; DB 23; Length 484;

Best Local Similarity 100.0%; Pred. No. 8.7e-222; Mismatches 0; Gaps 0;

Matches 424; Conservative 0; Indels 0; Gaps 0;

QY 1 MSYKAAGEBYKADCPGPNPAPTSMNGPDATAEEDFVPTVYTSAAKGDYKLVRF 60
 DB 48 MSYKAAGEBYKADCPGPNPAPTSMNGPDATAEEDFVPTVYTSAAKGDYKLVRF 107
 QY 61 GSSKIDKELINRIRATGQRPHFLRGIFPSHDMQVLDAYENKKPFYLYTGRGSSSE 120
 DB 108 GSSKIDKELINRIRATGQRPHFLRGIFPSHDMQVLDAYENKKPFYLYTGRGSSSE 167
 QY 121 AMVHGHILPFTFTWLODVPRVPLVIQMTDEKLMDLTLDOYGVAVENAKDIIACGF 180
 DB 168 AMVHGHILPFTFTWLODVPRVPLVIQMTDEKLMDLTLDOYGVAVENAKDIIACGF 227
 QY 181 DINKTFIFSDLDVYMGSSGFYKAVKIQKHTFNQVKGIFGFTSDICIGISFPALDAP 240
 DB 228 DINKTFIFSDLDVYMGSSGFYKAVKIQKHTFNQVKGIFGFTSDICIGISFPALDAP 287
 QY 241 SFSNSFQIFRDRITDIOCLIPCAIDODPYFPMTRDVAPRIGYKPPALLHSTFPALQGAQ 300

DB 288 SFSNSFQIFRDRITDIOCLIPCAIDODPYFPMTRDVAPRIGYKPPALLHSTFPALQGAQ 347
 QY 301 TKMSASDPNNSITLTPTAKQIKTKVKNKHPASGGARDTIEEHROGNCDDVDSFMYLTFFL 360
 DB 348 TKMSASDPNNSITLTPTAKQIKTKVKNKHPASGGARDTIEEHROGNCDDVDSFMYLTFFL 407
 QY 361 EDDDKLEQIRKQVTSAMLTGELKKALIEVLOPLIAEHOARRKEVDEIVKEFMTPRKLS 420
 DB 408 EDDDKLEQIRKQVTSAMLTGELKKALIEVLOPLIAEHOARRKEVDEIVKEFMTPRKLS 467
 QY 421 FDFQ 424
 DB 468 FDFQ 471

RESULT 5
 AAB58220
 ID AAB58220 standard; Protein; 475 AA.

AC AAB58220;

DT 14-MAR-2001 (first entry)

XX Lung cancer associated polypeptide sequence SEQ ID 558.

XX Human; lung cancer associated protein; neuroprotective; cytoskeletal;
 XX cardioprotective; immunomodulatory; muscular active; vulnereary;
 XX gastrointestinal; nephrotoxic; antineoplastic; gynecological;
 XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 XX proliferative disorder; wound healing; infectious disease.

OS Homo sapiens.

XX WO200055180-A2.

XX 21-SEP-2000.

PF 08-MAR-2000; 2000MO-US05918.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

PI Ruben SM;

DR WPI; 2000-587514/55.

DR N-PSDB; AAF18096.

PT Lung cancer associated gene sequences, referred to as lung cancer

PT antigens, useful for treatment, prevention, and diagnosis of disorders

PT such as lung cancer -

PS Claim 11; Page 1052-1053; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective, cytoskeletal, cardiovascular,
 CC immunomodulatory, muscular active general, vulnereary, gastrointestinal
 CC general, nephrotoxic, antineoplastic, gynecological, or antibacterial
 CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer.
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders. The proteins may also be used in the treatment of wounds and
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
 CC peptide AAB58549 are used in the course of the invention for the
 CC identification and characterisation of the polynucleotide and protein
 CC sequences.

XX SQ Sequence 475 AA;
 Query Match 99.3%; Score 2231; DB 21; Length 475;
 Best Local Similarity 99.3%; Pred. No. 2,9e-220;
 Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSYKAAAGEDYKADCPGPNAPTSNHPDPAEEDFVDPWTVQTSASAGIDYDKLIYRF 60
 DB 52 MSYKAAAGEDYKADCPGPNAPTSNHPDPAEEDFVDPWTVQTSASAGIDYDKLIYRF 111
 QY 61 GSSKIDKELINRIERATGQRPHHFLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSE 120
 DB 112 GSSKIDKELINRIERATGQRPHHFLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSE 171
 QY 121 AMHVGLIPIFTKMLQDVFNVLVYQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGF 180
 DB 172 AMHVGLIPIFTKMLQDVFNVLVYQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGF 231
 QY 181 DINKTFIFSDLDYMGSSGFYKAVVKIQKHVTFNOYKIGFPTSDCIGKISFPALQAA 240
 DB 232 DINKTFIFSDLDYMGSSGFYKAVVKIQKHVTFNOYKIGFPTSDCIGKISFPALQAA 291
 QY 241 SFSNSFPQIFRDRDIOCLIPCAIDODPYFRMTRDVAIRIGYKPKALLHSTFPALQAA 300
 DB 292 SFSNSFPQIFRDRDIOCLIPCAIDODPYFRMTRDVAIRIGYKPKALLHSTFPALQAA 351
 QY 301 TKMSASDPNSSIFLTDTAQIKTKVNHAFSGGRDTIEHRQFGNCDDVVSFMYLTFFL 360
 DB 352 TKMSASDPNSSIFLTDTAQIKTKVNHAFSGGRDTIEHRQFGNCDDVVSFMYLTFFL 411
 QY 361 EDDDKLEQIRKDYTSGLMTELKALIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLS 420
 DB 412 EDDDKLEQIRKDYTSGLMTELKALIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLS 471
 QY 421 PDPQ 424
 DB 472 PDPQ 475

RESULT 6
 AAY05372
 ID AAY05372 standard; Protein; 471 AA.
 AC AAY05372;
 DT 30-JUN-1999 (first entry)
 DE Human HCMV inducible gene protein, SEQ ID NO 12.
 XX
 XX HCMV inducible gene; c1g; human; human cytomegalovirus; interferon;
 KM anti-viral therapy; anti-HCMV therapy; detection; diagnosis;
 KM drug screening.
 XX
 XX Homo sapiens.
 OS
 XX
 XX MO9913075-A2.
 PN
 XX
 XX 18-MAR-1999.
 PD
 XX
 XX 08-SEP-1998; 98WO-US18638.
 PF
 XX
 XX 22-SEP-1997; 97US-0059725.
 PR
 XX 08-SEP-1997; 97US-0058180.
 PA (UYPR-) UNITV PRINCETON.
 XX
 XX Cong J, Schenk T, Zhu H;
 PI
 XX WPI; 1999-243729/20.
 DR
 XX N-PSDB; AAX33942.
 DR
 XX
 XX New isolated human genes
 PT

XX PS Claim 3; Page 112-114; 184pp; English.
 XX
 CC This sequence is encoded by a human gene of the invention, and is induced
 CC to express by both HCMV and interferon (IFN), designated HCMV-inducible
 CC genes (c1g or c1gs). The invention also relates to genes that are
 CC repressed in the presence of HCMV infection, designated HCMV-repressible
 CC genes (crg or crgs). The products can be used to obtain agents which can
 CC be used for anti-viral therapy, particularly anti-HCMV therapy. They can
 CC also be used for the development of drugs that would allow for higher
 CC dosage IFN treatments without the concomitant toxicity normally
 CC associated with administering high levels of IFN. The products can also
 CC be used for detection, diagnosis and drug screening.
 CC
 XX SQ Sequence 471 AA;
 Query Match 99.1%; Score 2226; DB 20; Length 471;
 Best Local Similarity 99.3%; Pred. No. 9.5e-220;
 Matches 421; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSYKAAAGEDYKADCPGPNAPTSNHPDPAEEDFVDPWTVQTSASAGIDYDKLIYRF 60
 DB 48 MSYKAAAGEDYKADCPGPNAPTSNHPDPAEEDFVDPWTVQTSASAGIDYDKLIYRF 107
 QY 61 GSSKIDKELINRIERATGQRPHHFLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSE 120
 DB 108 GSSKIDKELINRIERATGQRPHHFLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSE 167
 QY 121 AMHVGLIPIFTKMLQDVFNVLVYQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGF 180
 DB 168 AMHVGLIPIFTKMLQDVFNVLVYQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGF 227
 QY 181 DINKTFIFSDLDYMGSSGFYKAVVKIQKHVTFNOYKIGFPTSDCIGKISFPALQAA 240
 DB 228 DINKTFIFSDLDYMGSSGFYKAVVKIQKHVTFNOYKIGFPTSDCIGKISFPALQAA 287
 QY 241 SFSNSFPQIFRDRDIOCLIPCAIDODPYFRMTRDVAIRIGYKPKALLHSTFPALQAA 300
 DB 288 SFSNSFPQIFRDRDIOCLIPCAIDODPYFRMTRDVAIRIGYKPKALLHSTFPALQAA 347
 QY 301 TKMSASDPNSSIFLTDTAQIKTKVNHAFSGGRDTIEHRQFGNCDDVVSFMYLTFFL 360
 DB 348 TKMSASDPNSSIFLTDTAQIKTKVNHAFSGGRDTIEHRQFGNCDDVVSFMYLTFFL 407
 QY 361 EDDDKLEQIRKDYTSGLMTELKALIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLS 420
 DB 408 EDDDKLEQIRKDYTSGLMTELKALIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLS 467
 QY 421 PDPQ 424
 DB 468 PDPQ 471

RESULT 7
 AAB47617
 ID AAB47617 standard; Protein; 415 AA.
 AC AAB47617;
 DT 07-JAN-2002 (first entry)
 DE Human supermini Typrs.
 XX
 XX Tyrosyl-tRNA synthetase; Typrs; Rossmann fold nucleotide binding domain;
 KM vascular endothelial cell function; burn; plastic surgery; abdomen;
 KM polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
 KM angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
 KM dermal ulcer; diabetic ulcer; endothelialization;
 KM tryptophanyl-tRNA synthetase; typrs; vascular graft surgery.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200174841-A1.
 PN


```

XX 11-OCT-2001.
PD
XX
XX 21-MAR-2001; 2001WO-US08966.
XX
XX 31-MAR-2000; 2000US-193471P.
XX
XX (SCRI ) SCRIpps RES INST.
XX
XX Schimmel P, Wakaugi K;
XX
XX WPI; 2001-626377/72.
XX
XX N-PSDB; AAA43604.
XX
XX
XX New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
XX
XX PT vascular endothelial function, in particular for regulating
XX
XX PT angiogenesis, tumor metastasis and treating myocardial infarction -
XX
XX PS Disclosure; Page 129-30; 150pp; English.
XX
XX The sequences given in AAA47615-18 show full length and truncated
XX
XX CC versions of tyrophanyl-tRNA synthetase (TyrRS). The truncated TyrRS of
XX
XX CC the invention comprises a Rossmann fold nucleotide binding domain, and
XX
XX CC is capable of regulating vascular endothelial cell function. It is of
XX
XX CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of
XX
XX CC full length TyrRS with polymorphonuclear leucocyte elastase. Truncated
XX
XX CC TyrRS is useful for regulating angiogenesis, tumor metastasis,
XX
XX CC enhancing angiogenesis to a graft, treating myocardial infarction,
XX
XX CC solid tumor, and a condition that would benefit from increased or
XX
XX CC decreased angiogenesis in a mammal, in particular humans. It is also
XX
XX CC useful in diagnosis and as a wound healing agent for treating wounds
XX
XX CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
XX
XX CC plastic surgery when reconstruction is required following a burn or
XX
XX CC for cosmetic purposes. It is particularly useful in the treatment of
XX
XX CC abdominal wounds where there is high risk of infection. Truncated TyrRS
XX
XX CC promotes endothelialization in vascular graft surgery and is used in
XX
XX CC conjunction with angiography to administer the angiogenic tRNA
XX
XX CC synthetase polypeptides or polynucleotides directly to the lumen and
XX
XX CC wall of the blood vessel.
XX
XX
XX Sequence 415 AA;
SQ
Query Match 94.2%; Score 2116; DB 22; Length 415;
Best Local Similarity 100.0%; Pred. No. 1,6e-208;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 24 SNHGPDATAEADFDVDPWVQTSSAKGIDVDKLVIRFGSSKIDKELINRIERATGQRP 83
XX
XX DB 2 SNHGPDATAEADFDVDPWVQTSSAKGIDVDKLVIRFGSSKIDKELINRIERATGQRP 61
XX
XX QY 84 FLARGIFFSHRDMNVQVDAVENKKEPFLYTCRGSSAMVGHILPIFTKWLQDVNVNP 143
XX
XX DB 62 FLARGIFFSHRDMNVQVDAVENKKEPFLYTCRGSSAMVGHILPIFTKWLQDVNVNP 121
XX
XX QY 144 LVIGMTDEKYLMDLTLDQAYGDAVENAKDIIACGFIDINKTIFISDLDMGSSGGFYKN 203
XX
XX DB 122 LVIGMTDEKYLMDLTLDQAYGDAVENAKDIIACGFIDINKTIFISDLDMGSSGGFYKN 181
XX
XX QY 204 VVKIQKVTNVQVKGIFGFTDSDCIGKISFPALQAAFSNSPQIRDRDRTDCLIPCA 263
XX
XX DB 182 VVKIQKVTNVQVKGIFGFTDSDCIGKISFPALQAAFSNSPQIRDRDRTDCLIPCA 241
XX
XX QY 264 IDDDPYRMRDVAPRIGYPRPALHSTFPALQAGATKMSASPNSSIFLTDKAKIKT 323
XX
XX DB 242 IDDDPYRMRDVAPRIGYPRPALHSTFPALQAGATKMSASPNSSIFLTDKAKIKT 301
XX
XX QY 324 KVNKAHSGSGRDTIEERQFGNCDVVSFWYLTFFLEDDDKLEQIRKDYTGAMLTGEL 383
XX
XX DB 302 KVNKAHSGSGRDTIEERQFGNCDVVSFWYLTFFLEDDDKLEQIRKDYTGAMLTGEL 361
XX
XX QY 384 KKAILEVLOPLIAEHQARKEVTDEIVKEFMTPRKLSPDFQ 424
XX
XX DB 362 KKAILEVLOPLIAEHQARKEVTDEIVKEFMTPRKLSPDFQ 402

```

```

RESULT 8
AAE13493
ID AAE13493 standard; Protein; 415 AA.
XX
XX AAE13493;
AC
XX
XX 12-FEB-2002 (first entry)
DT
XX
XX Human supermini tyrophanyl t-RNA synthetase in pET20B.
DE
XX
XX Human; tyrophanyl-tRNA synthetase; TyrRS; tyrosyl t-RNA synthetase;
XX
XX KW TyrRS; vascular endothelial cell function; angiogenesis; wound healing;
XX
XX KW re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;
XX
XX KW diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;
XX
XX KW angiography; gene therapy; tumour; inflammation; vascular permeability;
XX
XX KW rheumatoid arthritis; psoriasis; diabetic retinopathy.
XX
XX OS Homo sapiens.
XX
XX PN WO200175078-A1.
XX
XX PD 11-OCT-2001.
XX
XX PF 21-MAR-2001; 2001WO-US08975.
XX
XX PR 31-MAR-2000; 2000US-193471P.
XX
XX PA (SCRI ) SCRIpps RES INST.
XX
XX PI Schimmel P, Wakaugi K;
XX
XX DR WPI; 2002-010784/01.
XX
XX DR N-PSDB; AAD22484.
XX
XX PT Novel truncated tyrophanyl-tRNA synthetase polypeptides capable of
XX
XX PT regulating vascular endothelial cell function, preferably angiogenesis,
XX
XX PT is useful for treating solid tumor or suppressing tumor metastasis in
XX
XX PT mammal -
XX
XX PS Example 1; Page 129-130; 149pp; English.
XX
XX The patent discloses human aminoacyl tRNA synthetases, particularly
XX
XX CC truncated tyrophanyl-tRNA synthetases (TyrRS) comprising a Rossmann
XX
XX CC fold nucleotide binding domain and polynucleotides encoding them. The
XX
XX CC invention also relates to tyrosyl t-RNA synthetases (TyrRS). TyrRS
XX
XX CC sequences are useful for regulating vascular endothelial cell function,
XX
XX CC preferably angiogenesis. Angiogenic TyrRS sequences are useful as wound
XX
XX CC healing agents for re-vascularizing damaged tissues. They are useful for
XX
XX CC treating full-thickness wounds (e.g. dermal ulcers, including pressure
XX
XX CC sores, venous ulcers and diabetic ulcers), burns and injuries. TyrRS
XX
XX CC sequences can also be used in plastic surgery when reconstruction is
XX
XX CC required following a burn, other trauma, or even for cosmetic purposes.
XX
XX CC Angiogenic TyrRS is also used in association with surgery and following
XX
XX CC the repair of cuts, for promoting endothelialisation in vascular graft
XX
XX CC surgery and for repairing the damage of myocardial infarction and in
XX
XX CC conjunction with coronary bypass surgery by stimulating the growth of
XX
XX CC transplanted tissue. TyrRS is also used in conjunction with angiography.
XX
XX CC TyrRS DNAs are useful in gene therapy. TyrRS antibodies are used in
XX
XX CC immunoassays to detect the presence of tumours. They are also useful
XX
XX CC for blocking endogenous angiogenic activity and retard the growth of
XX
XX CC solid tumours. These antibodies may also be used to treat inflammation
XX
XX CC caused by increased vascular permeability, inhibiting the activity of
XX
XX CC TyrRS by antisense technology is useful for preventing further growth
XX
XX CC or even regress solid tumours, and for treating Rheumatoid arthritis,
XX
XX CC psoriasis, diabetic retinopathy, all of which are characterised by
XX
XX CC abnormal angiogenesis. The present sequence is human truncated
XX
XX CC tyrophanyl t-RNA synthetase (supermini TyrRS; residues 71-471 of
XX
XX CC full-length TyrRS protein) in pET20B.
XX
XX Sequence 415 AA;
SQ

```

Query Match 94.2%; Score 2116; DB 23; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1.6e-208;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 SNEGPDAEEDPVPDPVTVQSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPBH 83
 DB 2 SNEGPDAEEDPVPDPVTVQSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPBH 61
 QY 84 FLNRGIFFSHRDMNOVLDAVENKKPFLYTGRRPSSAMHVGHLIPFTKWLQDVFNVP 143
 DB 62 FLNRGIFFSHRDMNOVLDAVENKKPFLYTGRRPSSAMHVGHLIPFTKWLQDVFNVP 121
 QY 144 LVIQMTDEKYLKMDLTDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 203
 DB 122 LVIQMTDEKYLKMDLTDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 181
 QY 204 VKKIQKHTFNOYKGIFFGFTSDSCIGKISFPALIOAPSFNSFPQIFRDRDIOCLIPCA 263
 DB 182 VKKIQKHTFNOYKGIFFGFTSDSCIGKISFPALIOAPSFNSFPQIFRDRDIOCLIPCA 241
 QY 264 IDDDPYFRMTRDVAPRIIGYKPKPALHSTFPPALIOAGOTKMSASDPNSSIFLDTTAQIKT 323
 DB 242 IDDDPYFRMTRDVAPRIIGYKPKPALHSTFPPALIOAGOTKMSASDPNSSIFLDTTAQIKT 301
 QY 324 KVNKAHPSGGRDYTEHRQFGNCDDVSEFMYLTFLEDDDKLEQIRKDYTGAMLTGEL 383
 DB 302 KVNKAHPSGGRDYTEHRQFGNCDDVSEFMYLTFLEDDDKLEQIRKDYTGAMLTGEL 361
 QY 384 KKAILEVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFO 424
 DB 362 KKAILEVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFO 402

RESULT 9
 AAB47618
 ID AAB47618 standard; Protein; 392 AA.
 XX AAB47618;
 XX 07-JAN-2002 (first entry)
 DT Human inactive Tryps.
 DE
 XX Tyrosyl-tRNA synthetase; Tryps; Roessmann fold nucleotide binding domain;
 KW vascular endothelial cell function; burn; plastic surgery; abdomen;
 KW polymorphonuclear leucocyte elastase; angiogenesis; tumor metastasis;
 KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
 KW dermal ulcer; diabetic ulcer; endothelialization;
 KW tryptophanyl-tRNA synthetase; trypsin; vascular graft surgery.
 XX
 OS Homo sapiens.
 XX
 XX WO200174841-A1.
 PN
 XX 11-OCT-2001.
 PD
 XX 21-MAR-2001; 2001MO-US08966.
 PF
 XX 31-MAR-2000; 2000US-193471P.
 PR
 XX (SCRI) SCRIPPS RES INST.
 PA
 XX Schimmel P; Wakaugi K;
 PI
 XX WPI; 2001-626377/72.
 DR
 XX N-PSDB; AAH43605.
 XX
 XX New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
 PT vascular endothelial function, in particular for regulating
 PT angiogenesis, tumor metastasis and treating myocardial infarction
 XX
 PS Disclosure; Page 135-36; 150pp; English.
 XX

The sequences given in AAB47615-18 show full length and truncated
 CC versions of tryptophanyl-tRNA synthetase (Tryps). The truncated Tryps of
 CC the invention comprises a Roessmann fold nucleotide binding domain, and
 CC is capable of regulating vascular endothelial cell function. It is of
 CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of
 CC full length Tryps with polymorphonuclear leucocyte elastase. Truncated
 CC Tryps is useful for regulating angiogenesis, tumor metastasis,
 CC enhancing angiogenesis to a graft, treating myocardial infarction,
 CC solid tumor, and a condition that would benefit from increased or
 CC decreased angiogenesis in a mammal, in particular humans. It is also
 CC useful in diagnosis and as a wound healing agent for treating wounds
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
 CC plastic surgery when reconstruction is required following a burn or
 CC for cosmetic purposes. It is particularly useful in the treatment of
 CC abdominal wounds where there is high risk of infection. Truncated Tryps
 CC promotes endothelialization in vascular graft surgery and is used in
 CC conjunction with angiography to administer the angiogenic tRNA
 CC synthetase polypeptides or polynucleotides directly to the lumen and
 CC wall of the blood vessel.
 CC
 SQ Sequence 392 AA;
 Query Match 88.5%; Score 1988; DB 22; Length 392;
 Best Local Similarity 100.0%; Pred. No. 2.2e-195;
 Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPBHFLNRGIFFSHRDMNOVLDAVENK 106
 DB 2 SAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPBHFLNRGIFFSHRDMNOVLDAVENK 61
 QY 107 KEFYLTYGRRPSSAMHVGHLIPFTKWLQDVFNVPVLIQMTDEKYLKMDLTDQAYG 166
 DB 62 KEFYLTYGRRPSSAMHVGHLIPFTKWLQDVFNVPVLIQMTDEKYLKMDLTDQAYG 121
 QY 167 DAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKNVVKIQRHTFNOYKGIFFGFTSD 226
 DB 122 DAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKNVVKIQRHTFNOYKGIFFGFTSD 181
 QY 227 CIGKISFPALIOAPSFNSFPQIFRDRDIOCLIPCAIDQDPYFRMTRDVAPRIIGYKPKA 286
 DB 182 CIGKISFPALIOAPSFNSFPQIFRDRDIOCLIPCAIDQDPYFRMTRDVAPRIIGYKPKA 241
 QY 287 LHSFPPALIOAGOTKMSASDPNSSIFLDTTAQIKTKVNKAHPSGGRDYTEHRQFGN 346
 DB 242 LHSFPPALIOAGOTKMSASDPNSSIFLDTTAQIKTKVNKAHPSGGRDYTEHRQFGN 301
 QY 347 CDDVSEFMYLTFLEDDDKLEQIRKDYTGAMLTGELKKAILEVLOPLIAEHQARRKEVT 406
 DB 302 CDDVSEFMYLTFLEDDDKLEQIRKDYTGAMLTGELKKAILEVLOPLIAEHQARRKEVT 361
 QY 407 DEIVKEFMTPRKLSFDFO 424
 DB 362 DEIVKEFMTPRKLSFDFO 379

RESULT 10
 AAE13494
 ID AAE13494 standard; Protein; 392 AA.
 XX AAE13494;
 XX 12-FEB-2002 (first entry)
 DT Human inactive tryptophanyl t-RNA synthetase in pET20B.
 DE
 XX Human; tryptophanyl-tRNA synthetase; Tryps; tyrosyl t-RNA synthetase;
 KW Tryps; vascular endothelial cell function; angiogenesis; wound healing;
 KW re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;
 KW diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;
 KW angiography; gene therapy; tumour; inflammation; vascular permeability;
 KW rheumatoid arthritis; psoriasis; diabetic retinopathy.
 XX
 OS Homo sapiens.
 XX

XX
PN WO200175078-A1.
XX
PD 11-OCT-2001.
XX
PF 21-MAR-2001; 2001WO-US08975.
XX
PR 31-MAR-2000; 2000US-193471P.
XX
PA (SCRI) SCRIIPS RES INST.
XX
PI Schimmel P, Wakaugi K;
XX
DR WPI; 2002-010784/01.
XX
DR N-PSDB; AAD22485.
XX
PT Novel truncated cryptophanyl-tRNA synthetase polypeptides capable of
PT regulating vascular endothelial cell function, preferably angiogenesis,
PT is useful for treating solid tumor or suppressing tumor metastasis in
PT mammal
XX
PS Disclosure, Page 135-136, 149pp; English.
XX
XX The patent discloses human aminoacyl tRNA synthetases, particularly
CC truncated cryptophanyl-tRNA synthetases (Trprs) comprising a Rossmann
CC fold nucleotide binding domain and polynucleotides encoding them. The
CC invention also relates to tyrosyl t-RNA synthetases (TyrRS). Trprs
CC sequences are useful for regulating vascular endothelial cell function,
CC preferably angiogenesis. Angiogenic Trprs sequences are useful as wound
CC healing agents for re-vascularizing damaged tissues. They are useful for
CC treating full-thickness wounds (e.g. dermal ulcers, including pressure
CC sores, venous ulcers and diabetic ulcers), burns and injuries. Trprs
CC sequences can also be used in plastic surgery when reconstruction is
CC required following a burn, other trauma, or even for cosmetic purposes.
CC Angiogenic Trprs is also used in association with surgery and following
CC the repair of cuts, for promoting endothelialisation in vascular graft
CC surgery and for repairing the damage of myocardial infarction and in
CC conjunction with coronary bypass surgery by stimulating the growth of
CC transplanted tissue. Trprs is also used in conjunction with angiography.
CC Trprs DNAs are useful in gene therapy. Trprs antibodies are used in
CC immunoassays to detect the presence of tumors. They are also useful
CC for blocking endogenous angiogenic activity and retard the growth of
CC solid tumours. These antibodies may also be used to treat inflammation
CC caused by increased vascular permeability. Inhibiting the activity of
CC Trprs by antisense technology is useful for preventing further growth
CC or even regress solid tumours, and for treating rheumatoid arthritis,
CC psoriasis, diabetic retinopathy, all of which are characterised by
CC abnormal angiogenesis. The present sequence is human inactive
CC cryptophanyl t-RNA synthetase (Trprs) in pET20B.
XX
SQ Sequence 392 AA;
Query Match 88.5%; Score 1988; DB 23; Length 392;
Best Local Similarity 100.0%; Pred. No. 2,2e-195;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 SAKSIDYDKLIVRGSSKIDKELINRIERATGQRPHPLRGIFSRDMNQVDAVENK 106
DB 2 SAKSIDYDKLIVRGSSKIDKELINRIERATGQRPHPLRGIFSRDMNQVDAVENK 61
QY 107 KPFLYLTGRGSSSEAMHVGHLIPITFTKWLQDVNVNVLVIMTDEKYLKMDLTLDQAYG 166
DB 62 KPFLYLTGRGSSSEAMHVGHLIPITFTKWLQDVNVNVLVIMTDEKYLKMDLTLDQAYG 121
QY 167 DAVENADIIACGCDINKTPIFSDLDMGSSGGYKAVVKIQRKVTNOKKGIIGFTDSD 226
DB 122 DAVENADIIACGCDINKTPIFSDLDMGSSGGYKAVVKIQRKVTNOKKGIIGFTDSD 181
QY 227 CIGKISPPAIQAAFSFNSFPQIRDRDTIOCLPCAIIDOPYFRMRDVAAPRGYKPA 286
DB 182 CIGKISPPAIQAAFSFNSFPQIRDRDTIOCLPCAIIDOPYFRMRDVAAPRGYKPA 241
QY 287 LHSSTFPALQAGQTKMSASDPNSSIFLDTAKQIKTKVNGAFSGGRDTIEHRQFGN 346

DB 242 LHSSTFPALQAGQTKMSASDPNSSIFLDTAKQIKTKVNGAFSGGRDTIEHRQFGN 301
QY 347 CDVDVSEFMILTFFLEDDDKLEQIRKDYTSGAMLTGELKALLIEVLOPLIAEHARKKEYT 406
DB 302 CDVDVSEFMILTFFLEDDDKLEQIRKDYTSGAMLTGELKALLIEVLOPLIAEHARKKEYT 361
QY 407 DEIVKEFMTPRKLSFDQ 424
DB 362 DEIVKEFMTPRKLSFDQ 379
RESULT 11
ABB64621
ID ABB64621 standard; Protein; 430 AA.
XX
XX ABB64621;
AC
XX 26-MAR-2002 (first entry)
DT
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 20655.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
PN WO200171042-A2.
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
PI Venter JC, Adams M, Li PMD, Myers EW;
XX
XX WPI; 2001-65860/75.
DR N-PSDB; ABL08724.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions
XX
XX Disclosure; SEQ ID NO 20655; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS7072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at fcp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 430 AA;
Query Match 61.4%; Score 1378; DB 22; Length 430;
Best Local Similarity 61.4%; Pred. No. 1.3e-132;
Matches 262; Conservative 64; Mismatches 85; Indels 16; Gaps 3;
QY 10 DYKADCPGNGAPTSNMGDA-----TEAE-----EDFVDEMTVQTSAGKIDYD 54
DB 3 DTEKTEVEGEALTLNGKDPABEVEGTDAQAQEGATAPEDVVDPMNVASSDAGVDYD 62
QY 55 KLIVRGSSKIDKELINRIERATGQRPHPLRGIFSRDMNQVDAVENKPFILYTG 114

Db 63 KLIRKRGSSKIDELIARFEKITGKPAHPIRRGWFESHRLDTLITLREQKPFYLYTG 122
 QY 115 RGPSSAMVGHILIPFIPTKMLQDVFNVPVLTQMTDEKYLWKDILTDQAYGDAVENAD 174
 Db 123 RGPSSGLVHGLVFPIMTKMLQETPDVPLVLTQMTDEKYLWKDILTDQAYGDAVENAD 182
 QY 175 IIAAGFDINKTIFESDLIDYMGSSGFYKNNVKIQKHVTFNQVGIIFGFTSDICIGKISFP 234
 Db 183 IYALIGDVNKTIFENLFEVKGCPAMYQNIIRIQKCVTFNQVGIIFGFTSDICIGKISFP 242
 QY 235 AIOAAPSFNSFPQIFRDTDIOCLIPCAIDDDPYFRMTDVAIRIGYKPKALLHSTFEP 294
 Db 243 AAOAAPALSTFPIFGNR-KVHCILPCALDDPYFRMTDVAIRIGYKPKALLHSTFEP 301
 QY 295 AIOGAQTMSASDPNSSIFLTDTAKQIKTKVNHAFSGGRTIEHRQPGNCDDVVSFM 354
 Db 302 AIOGAQTMSASDPNSSIFLTDTAKQIKTKVNHAFSGGRTIEHRQPGNCDDVVSFM 361
 QY 355 YLTFPLEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLAEHQARRKVTDEIVKEFM 414
 Db 362 LKTFPLEDDAKLEEVAVVASKGEMLTGEIKGLAVETLTPVEQHQAARKLITDEVLDKTF 421
 QY 415 TPRKLSF 421
 Db 422 ELRPLKF 428

RESULT 12

ABB67203
 ID ABB67203 standard; Protein; 430 AA.

AC ABB67203;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 28401.

DE Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR N-PSDB; ABL11306.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

PS Disclosure; SEQ ID NO 28401; 21bp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB11840-AB116175) and the encoded proteins (AB11840-AB116175). The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

QY Sequence 430 AA;

Query Match 61.4%; Score 1378; DB 22; Length 430;
 Best Local Similarity 61.4%; Pred. No. 1.3e-132;
 Matches 262; Conservative 64; Mismatches 85; Indels 16; Gaps 3;

QY 10 DYKADCPGNGPAPTSNHGPD-----TEAE-----EDFVDPWTQTSANGIDVD 54
 Db 3 DTKETVEGVEALTLTGKEDAPVEGTDAQOEQATAPTEEDVDVPMNVASNDAGVDYD 62
 QY 55 KLIVRGSSEKIDKELINRTERATGQRPHELRAGTIFPSHRDMQVLDVAENKKPFYLYTG 114
 Db 63 KLIRKRGSSKIDELIARFEKITGKPAHPIRRGWFESHRLDTLITLREQKPFYLYTG 122
 QY 115 RGPSSAMVGHILIPFIPTKMLQDVFNVPVLTQMTDEKYLWKDILTDQAYGDAVENAD 174
 Db 123 RGPSSGLVHGLVFPIMTKMLQETPDVPLVLTQMTDEKYLWKDILTDQAYGDAVENAD 182
 QY 175 IIAAGFDINKTIFESDLIDYMGSSGFYKNNVKIQKHVTFNQVGIIFGFTSDICIGKISFP 234
 Db 183 IYALIGDVNKTIFENLFEVKGCPAMYQNIIRIQKCVTFNQVGIIFGFTSDICIGKISFP 242
 QY 235 AIOAAPSFNSFPQIFRDTDIOCLIPCAIDDDPYFRMTDVAIRIGYKPKALLHSTFEP 294
 Db 243 AAOAAPALSTFPIFGNR-KVHCILPCALDDPYFRMTDVAIRIGYKPKALLHSTFEP 301
 QY 295 AIOGAQTMSASDPNSSIFLTDTAKQIKTKVNHAFSGGRTIEHRQPGNCDDVVSFM 354
 Db 302 AIOGAQTMSASDPNSSIFLTDTAKQIKTKVNHAFSGGRTIEHRQPGNCDDVVSFM 361
 QY 355 YLTFPLEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLAEHQARRKVTDEIVKEFM 414
 Db 362 LKTFPLEDDAKLEEVAVVASKGEMLTGEIKGLAVETLTPVEQHQAARKLITDEVLDKTF 421
 QY 415 TPRKLSF 421
 Db 422 ELRPLKF 428

RESULT 13

AAG23698
 ID AAG23698 standard; Protein; 402 AA.

AC AAG23698;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 27101.

DE Protein identification; signal transduction pathway; metabolic pathway;

KM hybridization assay; genetic mapping; gene expression control; promoter;

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.

```

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160747.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 58.1%; Score 1304; DB 21; Length 402;
Best Local Similarity 60.3%; Pred. No. 4.9e-125;
Matches 240; Conservative 67; Mismatches 87; Indels 4; Gaps 2;

```

```

QY 29 DATEAE--EDFVPMPTVQTSAGIDYDKLIVFGSSKIDKELINRIERATGQRPHFPR 86
DB 7 DEREASSEVQVNPWEVSADGCKIDYDKLIDFGCORDESIDYQRLTSQPHVPR 66
QY 87 RGIFSFHDMQVLAIDENKPPYLYTGRGSSSEAMVGLIPFTKMLQDVFNPLVI 146
DB 67 RSVFPAHRDENEILDAVERGDKFYLYTGRGSSSEALHGLIPFTKYLQEAFFKPLVI 126
QY 147 QMTDDKYLKMLTLOAQYDAVENAKDIIACGPDINKPTIFSDLLYMGSSGFYQNVK 206
DB 127 QLTDDKSKYWKNSIVSESRLEAKENAKDIIACGPDYKTFIFSDFDYVG--GAFYQNVK 184
QY 207 IOKHVFNOVKYJFGFTSDICIKISFPALQAPFSNSPQIFRRTDIOCLIPCAIO 266
DB 185 VGCYVLRNMGKIFGSGEPRIKLSFPYQAVPSPSPSPHIFPGKDNIRCLIPALIO 244
QY 267 DPFYRMTDVAERIGYKPKPALHSTFFPALQGAQTKMSADPNSSIFLDTAKQIKTKV 326
DB 245 DPFYRMTDVAERIGYKPKPALHSTFFPALQGENKMSADPNSSALYVTDASAKDINKIN 304
QY 327 KHAFGSGRPTIEHROFGNCVDVSMYITFLEDDDKLQIRKDYTSGAMLTGLKXA 386
DB 305 RAHFSGGQDSIEHREIGANLEVDIPKYLISFLEDDSELEHKEYGGRMLTGEVKR 364
QY 387 LIEVLOPLIAEHOARRKEVTEIYKEFMTPRKLSFPO 424
DB 365 LIEVLEIYKRRARAATVDENVDAFMAVRPLPSKFE 402

```

```

RESULT 14
AAG23697 standard; Protein; 426 AA.
AAG23697;

```

```

XX 17-OCT-2000 (first entry)

```

```

DE Arabidopsis thaliana protein fragment SEQ ID NO: 27100.

```

```

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

```

```

OS Arabidopsis thaliana.

```

```

XX PN EP1033405-A2.

```

```

XX PD 06-SEP-2000.

```

```

XX PF 25-FEB-2000; 2000EP-0301439.

```

```

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136722.
PR 01-JUN-1999; 99US-0137282.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0138119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.

```


DB 389 LTEVLTEIVEKRRARAATDEKMDAFMAVRPLPSKFE 426

RESULT 15

ID AAB66931 standard; Protein; 424 AA.

AC AAB66931;

DT 12-APR-2001 (first entry)

DE Tryptophanyl-tRNA synthetase.

KW Tryptophanyl-tRNA synthetase; enzyme.

OS Candida albicans.

PN US6174713-B1.

PD 16-JAN-2001.

PF 16-JUN-1997; 97US-0876685.

PR 16-JUN-1997; 97US-0876685.

PA (CUBI-) CUBIST PHARM INC.

PI Shen X, Houman F;

XX WPI; 2001-201806/20.

DR N-PSDB; AAF55855.

PT New nucleic acids encoding Candida cytoplasmic tryptophanyl-tRNA

PT (cRNA) synthetases, useful for producing recombinant cRNA synthetases

PS and detecting inhibitor of Candida cRNA synthetase function

CC The present sequence is Candida albicans cytoplasmic tryptophanyl-tRNA

CC (cRNA) synthetase; cRNA synthetase is useful for producing recombinant

CC cRNA synthetases and detecting inhibitors of cRNA synthetase.

SQ Sequence 424 AA;

Query Match 54.3%; Score 1218.5; DB 22; Length 424;

Best Local Similarity 57.1%; Pred. No. 3.3e-116; Mismatches 92; Indels 7; Gaps 3;

Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

DB 31 TEAEEDFVPMWTVQ---TSAGKIDYDKLIVFGSSKIDKELINRIERATGGRPHPLR 86

DB 13 TESEBQKITPWEYEGAVNDCKSGIDYDKLISQGTGKHITETLERFKOYTGEEHPFLK 72

QY 87 RGIFFSHRDMDNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFIPTKWLQDVENVPLVI 146

DB 73 RGVFFSQRDLDRLIDLYEHGEPEFLYTGRGSSSDSMHLSHMVVFITKWLQEVFVPLVI 132

QY 147 QMTDDDEKYLAK-DLITDQAVGDVAVENAKDIIACFPDINKFTIFSOLDYMGSSGFYKAV 205

DB 133 ELTDDDEKFLFKQDLTIDVKGFAENAKDIIAGFNPNFTIFSOLQYMG--GAFYENVV 190

QY 206 KIQKHTVFNQVKGIFGFTSDSCIGKISFPALQAPFSNSFPQIFRDRDIOCLIPCAID 265

DB 191 KTSRQITTSATAVAFSTDSDCIGKIHFAISIQATATFSSFPVGLPKPTKPLIIPCAID 250

QY 266 QDPYFMTDVAARIVPKPALIHSTFPALQAGQTKMSADPNSSIFLDTAKQIKTV 325

DB 251 ODPYFVVCDDVADKLIFTKPALIHAKFPALQASTKMSADTTTISIFMGDTAKQIQKKI 310

QY 326 NKHAFSGGSDTJEHROFGNCDVDSFWMTLFELEDDDKLQIRKDYTSGLMTGELKK 385

DB 311 NKAFSGGSRATAEHRELGNPEVDVAFQYLSFSYDEKLAQLEQYRKGETLSEMKK 370

DB 371 ECITVLQEFVSAYQERRSKVDQDVVEKFMKPKLVF 406

Search completed: July 10, 2003, 12:30:11
Job time : 35.92 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:29:05 ; Search time 12.8485 Seconds
(without alignments)
970.956 Million cell updates/sec

Title: US-09-813-718-10_COPY_48_471
Perfect score: 2246
Sequence: 1 MSYKAAGBDYKADCPPGNP.....VTDEIVKEFMTPRKLSDFDQ 424

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1218.5	54.3	424	US-08-876-885-26	Sequence 26, Appl
2	185.5	8.3	341	US-08-928-100-2	Sequence 2, Appl
3	185.5	8.3	341	US-09-492-581-2	Sequence 2, Appl
4	185.5	8.3	341	US-09-425-666-2	Sequence 2, Appl
5	164.5	7.3	409	US-08-743-130A-39	Sequence 39, Appl
6	162.5	7.2	409	US-08-743-130A-2	Sequence 2, Appl
7	132.5	5.9	388	US-08-705-868-4	Sequence 4, Appl
8	132.5	5.9	388	US-09-123-615-4	Sequence 4, Appl
9	113.5	5.1	418	US-08-855-910-11	Sequence 11, Appl
10	109	4.9	377	US-09-352-990-28	Sequence 28, Appl
11	108	4.8	197	US-08-823-867-2	Sequence 2, Appl
12	108	4.8	197	US-08-928-100-4	Sequence 4, Appl
13	108	4.8	197	US-09-183-134-2	Sequence 2, Appl
14	108	4.8	197	US-09-492-581-4	Sequence 4, Appl
15	108	4.8	197	US-09-425-666-4	Sequence 4, Appl
16	104.5	4.7	370	US-08-415-593-45	Sequence 45, Appl
17	100.5	4.5	427	US-09-134-001C-5141	Sequence 5141, Ap
18	95	4.2	418	US-08-844-054-2	Sequence 2, Appl
19	95	4.2	418	US-09-347-333-2	Sequence 2, Appl
20	95	4.2	464	US-09-134-001C-4701	Sequence 4701, Ap
21	94.5	4.2	877	US-08-907-166-8	Sequence 8, Appl
22	93	4.1	607	US-09-134-001C-2994	Sequence 2994, Ap
23	90.5	4.0	898	US-08-465-995A-4	Sequence 4, Appl
24	90.5	4.0	898	US-08-465-995A-4	Sequence 4, Appl
25	90.5	4.0	898	US-08-465-995A-4	Sequence 4, Appl
26	90.5	4.0	920	US-08-101-593-4	Sequence 4, Appl
27	89.5	4.0	344	US-09-393-554-2	Sequence 2, Appl

28	89.5	4.0	898	US-08-465-995A-2	Sequence 2, Appl
29	89.5	4.0	898	US-08-465-995A-2	Sequence 2, Appl
30	89.5	4.0	898	US-08-465-995A-2	Sequence 2, Appl
31	89.5	4.0	920	US-08-101-593-2	Sequence 2, Appl
32	89	4.0	719	US-08-082-849B-31	Sequence 31, Appl
33	89	4.0	719	PCT-US94-01624-31	Sequence 31, Appl
34	88.5	3.9	501	US-09-157-257-8	Sequence 8, Appl
35	88	3.9	606	US-08-883-534-3	Sequence 3, Appl
36	88	3.9	606	US-09-204-764-3	Sequence 3, Appl
37	87.5	3.9	855	US-08-890-865A-10	Sequence 10, Appl
38	87	3.9	428	US-08-331-625A-43	Sequence 43, Appl
39	87	3.9	502	US-09-494-151-43	Sequence 43, Appl
40	87	3.9	428	US-09-134-001C-4511	Sequence 4511, Ap
41	87	3.9	1452	US-08-331-625A-2	Sequence 2, Appl
42	87	3.9	1452	US-09-494-151-2	Sequence 2, Appl
43	87	3.9	1452	PCT-US93-04384-18	Sequence 18, Appl
44	87	3.9	1452	PCT-US93-04692-2	Sequence 2, Appl
45	86.5	3.9	849	US-09-157-257-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-876-885-26

Sequence 26, Application US/08876885

Patent No. 6174713

GENERAL INFORMATION:

APPLICANT: Shen, Xiaoyu

APPLICANT: Houman, Fariba

TITLE OF INVENTION: CANDIDA CYTOPLASMIC TRYPTOPANYL-TRNA

TITLE OF INVENTION: SYNTHETASE PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING

TITLE OF INVENTION: SAME

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: MA USA

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/876, 885

FILING DATE: 16-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Brook, David E.

REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: CPI97-02

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781) 861-6240

TELEFAX: (781) 861-9540

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 424 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-876-885-26

Query Match 54.3%; Score 1218.5; DB 4; Length 424;
Best Local Similarity 57.1%; Pred. No. 6.1e-127;
Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

QY 31 TEAEEDFDPFWQ---TSSAKGIDYDKLIVRGSSKIDKELNIRIRATGQRPHHFLR 86
DB 13 TEESOKITPWEVEGAVVDGKSMGIDYDKLISOFKTHITESTLERFQOVGGEHPPLK 72

```

Qy 8 RGFESIRHNNQVLDAYENKKEPFYLTGSGPSSAMHGHILPFIETMILQDVNVPVLVI 146
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 73 RGVFPSSGRDRLRLDLYEHGEPPEFLYTGKSGPSSMHLGHMVPFIETMILQDEVDPVLVI 1322
Qy 147 QMTDEKYLTK -DLTLDQAYGDAVENAKOITACGDPDKITPFIPSLDYMGMSSEFYKNV 2050
   ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 133 ELTLDDEKPEFKOGLTLDVKGFAEENAKOITAVGNPNENTFPIPSLOQTMG--GAFYENNV 1900
Qy 206 KIOKHVFNQVKIIFGFTSDSCIGKISPAIOAESPNSFPOLTRDUTDIOCLIPCAID 2655
   ::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 191 RTSRQITTSYAKAVVFGFTSDSCIGKIHFASIGIATAFPSSFPDVLGLPEPTKPLICPAID 2500
Qy 266 ODYFEMTTPVAVRIGYPPKRALHSTFPALQAGATKASADPNNSIFLUTPAOIKTV 3250
   ||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 251 QDPYFRCVAVAKLFTPKPALHAAFPALQAGATKASASTTTSIPMGTAIOIPKTI 3100
Qy 326 NKHAFSGGRITIEIENHQFGNCQVDVSPMYILTFPEDDDKLEOIRKDYTSGAMLTGELK 3850
   ||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 311 NKYAFSGRATAEHEHLELGNPEVDVAFOYLSFSPSYDEKLAQLAQOGRYKGEILSGEMKK 3700
Qy 386 ALLEVLQPLAENHQAARKVEITDEIYVENMTTPKKSF 421
   ||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 371 ECIIVLOEFVSAQERRKSKVDQVVEKPEKPKLVAF 406

```

RESULT 2

US-08-928-100-2
; Sequence 2, Application US/08928100
; Patent No. 6046174

```

1 GENERAL INFORMATION:
2 APPLICANT: Gentley, Danile
3 APPLICANT: Greenwood, Claire
4 APPLICANT: Lawlor, Elizabeth
5 TITLE OF INVENTION: No. 6046174el trps
6 NUMBER OF SEQUENCES: 6
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: SmithKline Beecham Corporation
9 STREET: 709 Swedeland Road
10 CITY: King of Prussia
11 STATE: PA
12 COUNTRY: USA
13 ZIP: 19406-0939
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette
16 COMPUTER: IBM Compatible
17 OPERATING SYSTEM: DOS
18 SOFTWARE: FASTSEQ for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/928,100
21 FILING DATE: 12-SEP-1997
22 CLASSIFICATION: 424
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: 9619072.3
25 FILING DATE: 12-SEP-1996
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Gilmel, Edward R
28 REGISTRATION NUMBER: 38,891
29 REFERENCE/DOCKET NUMBER: P31624-1
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 610-270-4478
32 TELEFAX: 610-270-5090
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978

```

Query Match	8.3%;	Score 185.5;	DB 3;	Length 341
Best Local Similarity	24.0%;	Pred. No. 5.4e-12;		

	Matches	81; Conservative	60; Mismatches	128; Indels	69; Gaps	15;
QY	107	KPEFLYLRGRGSSAMVGHILPIFTYQADVFNVPLVIQMTDEKY-LMKDLTLTDAY	165			
DB	3	KPIIL-TGDRPTK-LHIGHYVGLKOR-----VLLQEDKDKMVEFLADQDAL	49			
QY	166	GDAVEN-----AKDIIAGCFDINKTEIF--SDLDWGMGSGEYKVV--KIQK	209			
DB	50	TDHAKDPQTVIESIGNVALDYIANGLDPNKSTIFIQSQIPELAEISMYNNLVSLAKLER	109			
QY	210	HVFENOVYKGIPTGTFSDICIGKISFPALQAA--PFSFNSFPQIFRDRDTIOCLIPCALDQD	267			
DB	110	NPTYKTBESQGFPEBSPITPTGLVYPIQAADITAFKANY-----VPTGTQK	156			
QY	268	PYFMRTRD-----VAPRIGYKPKALIHSTFPFALQAGQTKMSASDPNSSIF	313			
DB	157	PMIQTRIEIVRSPFNNAVNCVILVPEEGYIPENE--RAQRLTGLDQ--NAKMSKS-LNNGIV	212			
QY	314	LTDPAKQIKTKVNNKHAISGGRDITEHRQFQSGNCVDVSPMYLTFP--LEDDDKYLEQIRK	371			
DB	213	LADADLTIRKKVMSMYTDPDPIHIREVDPKIGEN---MVFFYLDVFGRPEDAQELTADNKE	268			
QY	372	DYTSGAMLTGELTKALIEVLQPLAEHQAAREVTDEI	409			
DB	269	RYQRGGLQDVTKKRYLLEIRELEQIPRKRIRKIAKDM	306			

RESULT 3

US-09-492-581-2
; Sequence 2, Application US/09492581
; Patent No. 6346409

```

1 GENERAL INFORMATION:
2 APPLICANT: Gentry, Danile
3 APPLICANT: Greenwood, Claire
4 APPLICANT: Lawlor, Elizabeth
5 TITLE OF INVENTION: No. 6346409el trps
6 NUMBER OF SEQUENCES: 6
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: SmithKline Beecham Corporation
9 STREET: 709 Swedeland Road
10 CITY: King of Prussia
11 STATE: PA
12 COUNTRY: USA
13 ZIP: 19406-0939
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette
16 COMPUTER: IBM Compatible
17 OPERATING SYSTEM: DOS
18 SOFTWARE: FASTSEQ for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/09/492,581
21 FILING DATE:
22 CLASSIFICATION:
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US/08/928,100
25 FILING DATE: 12-SEP-1997
26 APPLICATION NUMBER: 9619072.3
27 FILING DATE: 12-SEP-1996
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Gimm1, Edward R
30 REGISTRATION NUMBER: 38,891
31 REFERENCE/DOCKET NUMBER: P31624-1
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: 610-270-4478
34 TELEFAX: 610-270-5090
35 TELEX:
36 INFORMATION FOR SEQ ID NO: 2:
37 SEQUENCE CHARACTERISTICS:
38 LENGTH: 341 amino acids
39 TYPE: amino acid
40 STRANDEDNESS: single
41 TOPOLOGY: linear
42 MOLECULE TYPE: protein

```

US-09-492-581-2

Query Match 8.3%; Score 185.5; DB 4; Length 341;

Best Local Similarity 24.0%; Pred. No. 5,4e-12;

Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 107 KEFYLYTGRGSPSEAMHGHILPIFTKMLQDVFNPLVIOQMTDDEKY-LMKDLTLDDQAY 165
 DB 3 KPIILL-TGDRPTGK-LHIGHVYSLKNR-----VLLQEDKXDMFVFLADQQL 49
 QY 166 GDAVEN-----ANDIACGFDINKTPIF--SDLDYGMSSGFKYXNV---KIQK 209
 DB 50 THAADPQIVTSIGNVALDYLAAGLDPNKSTFIQSQIPBELAELSMYMNIVSLARLER 109
 QY 210 HTFNOVKGIPTGTDSDCIKISFPAIOAA--PSFSNSFPQIFRDRTDIOCLIPCAIDOD 267
 DB 110 NPTVKTETISQKGFSGSIPFGFLVYPIAQAADITAFKANY-----VPVGTQK 156
 QY 268 PYFRMTRD-----VAPRIQYKPKALHSTFFPALQGAQTMSASDPNSSIF 313
 DB 157 PMIEQREIVRSFNNAVNCVLYVEPGIYPENE--RAGRLPGLDG-NAMKSKS-LNNGIY 212
 QY 314 LDTAKQIKTKVKNKAHFGSGRDTIEBHROFGNCDVVSFMYLTFE--LEDDKLEQIRK 371
 DB 213 LADDDADTLRKXKWSMTYDHPDIRVEDPGKIEGN---MVFHLYDVFGREPDAQEIADMKX 268
 QY 372 DYTSGAMLTGELKALIEVLOPIAHOARKEVTDEI 409
 DB 269 RYORGGLGDKTKRYLLETILERELGPIRRRIEFAKDM 306

RESULT 4

US-09-425-666-2

Sequence 2, Application US/09425666

Patent No. 6416976

GENERAL INFORMATION:

APPLICANT: Gentry, Danile

APPLICANT: Greenwood, Claire

APPLICANT: Lawlor, Elizabeth

TITLE OF INVENTION: No. 6416976el trps

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/425,666

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/928,100

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Gimm, Edward R

REGISTRATION NUMBER: 38,881

REFERENCE/DOCKET NUMBER: P31624-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 341 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-425-666-2

Query Match 8.3%; Score 185.5; DB 4; Length 341;

Best Local Similarity 24.0%; Pred. No. 5,4e-12;

Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 107 KEFYLYTGRGSPSEAMHGHILPIFTKMLQDVFNPLVIOQMTDDEKY-LMKDLTLDDQAY 165
 DB 3 KPIILL-TGDRPTGK-LHIGHVYSLKNR-----VLLQEDKXDMFVFLADQQL 49
 QY 166 GDAVEN-----ANDIACGFDINKTPIF--SDLDYGMSSGFKYXNV---KIQK 209
 DB 50 THAADPQIVTSIGNVALDYLAAGLDPNKSTFIQSQIPBELAELSMYMNIVSLARLER 109
 QY 210 HTFNOVKGIPTGTDSDCIKISFPAIOAA--PSFSNSFPQIFRDRTDIOCLIPCAIDOD 267
 DB 110 NPTVKTETISQKGFSGSIPFGFLVYPIAQAADITAFKANY-----VPVGTQK 156
 QY 268 PYFRMTRD-----VAPRIQYKPKALHSTFFPALQGAQTMSASDPNSSIF 313
 DB 157 PMIEQREIVRSFNNAVNCVLYVEPGIYPENE--RAGRLPGLDG-NAMKSKS-LNNGIY 212
 QY 314 LDTAKQIKTKVKNKAHFGSGRDTIEBHROFGNCDVVSFMYLTFE--LEDDKLEQIRK 371
 DB 213 LADDDADTLRKXKWSMTYDHPDIRVEDPGKIEGN---MVFHLYDVFGREPDAQEIADMKX 268
 QY 372 DYTSGAMLTGELKALIEVLOPIAHOARKEVTDEI 409
 DB 269 RYORGGLGDKTKRYLLETILERELGPIRRRIEFAKDM 306

RESULT 5

US-08-743-130A-39

Sequence 39, Application US/08743130A

Patent No. 5871987

GENERAL INFORMATION:

APPLICANT: Sassanfar, Mandana

APPLICANT: Gallant, Paul L.

APPLICANT: Shen, Xiaoyu

APPLICANT: Tao, Nianjun

APPLICANT: Tao, Jianshi

APPLICANT: Houman, Fariba

TITLE OF INVENTION: CANDIDA TYROSYL-ERNA SYNTHETASE

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: Massachusetts

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/743,130A

FILING DATE: 01-NOV-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Brook Esq., David E.

REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: CPI95-12

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240

TELEFAX: (617) 861-9540

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 409 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-743-130A-39

Query Match
Best Local Similarity 20.7%; Score 164.5; DB 2; Length 409;
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

7.3%; Score 164.5; DB 2; Length 409;
Best Local Similarity 20.7%; Pred. No. 1.6e-09;
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

99 VLDAYENK-KPFLYITGRGSSSEAMVGHILPFI-----FTKMLQDV-----F 140
27 IKVLEKENRPVIXYWGTAFTGKP-HCGYFVPMIKAHFLKAGCEVTLADLHAFLDNM 85
141 NVPLVQMTDDEKYLMDLTLDOAYGDVAVENAKDIIACGFDINKTIFSDLDVMGSSGF 200
86 KAPLEVVKYRAKYEFVVKAILKSNVPIERLKFVVGSSYQKGADV--MDLFKLSNIV 142
201 YKAVVK-----IOKHYTFNOVKGIFGFTSDICIGISFPALIOAPSPNSFPQIFRDRD 255
143 SQMDARAGADVVKQVANPLISGLI-----YPLMQA-----IDEEHLG 180
256 IQCLIPCAIDQDPYFRMTDVAPRIGYPKPALHSTFPALQAGQTMASDPNSIFLT 315
181 VDAQFG-GVDQRIKIFLAENLPSIGYKRAHLMNPMVPGI-GQGGKMSASDPNSKIDI 238
316 DTKAKQIKTKNKAIFSGG--RDT-----IEEHROGCGN 346
239 EEPKVVKKVNSAYCAPGELKDGLAFIEYVIQPIAELKTGVEGAFKLDIDREKYG- 297
347 CDVDSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKALI-----EVLQPLIAHQARR 402
298 ---PLSY-----DSIEQLKADIVDGKLAAPDLKSGVADKINLAPIRAEFESS- 343

QY 403 KEVTDEIVKEFMTPRK 418
DB 344 -----EEFOVAQK 351

RESULT 6
US-08-743-130A-2
Sequence 2, Application US/08743130A
Patent No. 5871987
GENERAL INFORMATION:
APPLICANT: Saseanfar, Mandana
APPLICANT: Gallant, Paul L.
APPLICANT: Shen, Xiaoyu
APPLICANT: Tao, Nianjun
APPLICANT: Houman, Fariba
TITLE OF INVENTION: CANDIDA TYROSYL-tRNA SYNTHETASE
TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743.130A
FILING DATE: 01-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook Esq., David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI95-12
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-743-130A-2

Query Match
Best Local Similarity 20.7%; Score 162.5; DB 2; Length 409;
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

7.2%; Score 162.5; DB 2; Length 409;
Best Local Similarity 20.7%; Pred. No. 2.7e-09;
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

99 VLDAYENK-KPFLYITGRGSSSEAMVGHILPFI-----FTKMLQDV-----F 140
27 IKVLEKENRPVIXYWGTAFTGKP-HCGYFVPMIKAHFLKAGCEVTLADLHAFLDNM 85
141 NVPLVQMTDDEKYLMDLTLDOAYGDVAVENAKDIIACGFDINKTIFSDLDVMGSSGF 200
86 KAPLEVVKYRAKYEFVVKAILKSNVPIERLKFVVGSSYQKGADV--MDLFKLSNIV 142
201 YKAVVK-----IOKHYTFNOVKGIFGFTSDICIGISFPALIOAPSPNSFPQIFRDRD 255
143 SQMDARAGADVVKQVANPLISGLI-----YPLMQA-----IDEEHLG 180
256 IQCLIPCAIDQDPYFRMTDVAPRIGYPKPALHSTFPALQAGQTMASDPNSIFLT 315
181 VDAQFG-GVDQRIKIFLAENLPSIGYKRAHLMNPMVPGI-GQGGKMSASDPNSKIDI 238
316 DTKAKQIKTKNKAIFSGG--RDT-----IEEHROGCGN 346
239 EEPKVVKKVNSAYCAPGELKDGLAFIEYVIQPIAELKTGVEGAFKLDIDREKYG- 297
347 CDVDSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKALI-----EVLQPLIAHQARR 402
298 ---PLSY-----DSIEQLKADIVDGKLAAPDLKSGVADKINLAPIRAEFESS- 343

QY 403 KEVTDEIVKEFMTPRK 418
DB 344 -----EEFOVAQK 351

RESULT 7
US-08-705-868-4
Sequence 4, Application US/08705868
Patent No. 5885798
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Coleman, Roger
APPLICANT: Au-Young, Janice
APPLICANT: Murty, Lynn E.
TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,868
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749

? REFERENCE/DOCKET NUMBER: PF-0117 US
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: 415-855-0555
 ? TELEFAX: 415-845-4166
 ? INFORMATION FOR SEQ ID NO: 4:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 388 amino acids
 ? TYPE: amino acid
 ? STRANDEDNESS: single
 ? TOPOLOGY: linear
 ? MOLECULE TYPE: peptide
 ? IMMEDIATE SOURCE:
 ? LIBRARY: GenBank
 ? CLONE: 1184699
 ?
 ?
 US-08-705-868-4

Query Match	5.9%	Score 132.5	DB 2	Length 388
Best Local Similarity	20.8%	Pred. No. 5.3e-06		
Matches 80; Conservative	58;	Mismatches 121;	Indels 125;	Gaps 20

Oy 114GRGGS-SEAMHVGHLIPFETKTKWLODVFNPLVIMONTDDE--KYLMKDLTLDD--AYGD 167
 Db 2*GDAPSPEEKLH-----LITRNLOEVLGEKELKEIKELKIYWGATGTGKPHVAYFV 54

Qy 168 AVENAKDITACGFEDNKTFPSDL-----DYGMSGGFKNVKVIQKH---VTF 213

Db 55 PMSKTADFLKAGCEV--TTFADLHA YLDNKKAPPELLERVSYENVVKMLSEIGVPL 112

QY	214	NYVKGIFG-----FTDSD-----	CIGKISFAIQQA	239
DB	113	EKLPIKIGDDYQLSKYEYTLDDYVRUSSVVTIHDSSKGA	GVVKKOVENHPLGSLGLYFELQ	171

QY 240 PPSNSNFPQIFRDRTDICCLPCALDDPPFRMTDVAPELGYRFPALHSTFFPALQGA 299
Db 172 --LDSEYLKV-----DAQF---GGIDQRKLTFAEKYLPALGYSTRVHLNMPMVPGLTG- 220

Qy 300 QTKMSASDNNSTFLDTAKQIKTVNKAFA-----SGGRDTIEH----- 340
:
:
:
Db 221 -SKSSSSSEESKIDLLDRKEDVKKLTK-AFCEPGNVNNGVLSPYKHVLPPLKSEFVIL 278
:
:
:

Cy 341 --RQFGNCVDVDSFMYLTFEFLLEDDKLEQIRKDYSGAALTGELKCALIEVLQPIAEH 398
 b 279 RDEKMGGN-----KTYATVVD-----LEKEDPAAEVHPGDIKNSVEALINKL----- 321

Qy	399	QARRKEVTDEIVKEFMT	--RKL	420
Qb	332	-----DPIREKENTPAL	KLA	337

RESULT 8
US-09-123-615-4

```

; Sequence 4, Application US/09125615
; Patent No. 6090377
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga

```

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive

CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304

```

1      ;
2      ;
3      ;
4      ;
5      ;
6      ;
7      ;
8      ;
9      ;
10     ;
11     ;
12     ;
13     ;
14     ;
15     ;
16     ;
17     ;
18     ;
19     ;
20     ;
21     ;
22     ;
23     ;
24     ;
25     ;
26     ;
27     ;
28     ;
29     ;
30     ;
31     ;
32     ;
33     ;
34     ;
35     ;
36     ;
37     ;
38     ;
39     ;
40     ;
41     ;
42     ;
43     ;
44     ;
45     ;
46     ;
47     ;
48     ;
49     ;
50     ;
51     ;
52     ;
53     ;
54     ;
55     ;
56     ;
57     ;
58     ;
59     ;
60     ;
61     ;
62     ;
63     ;
64     ;
65     ;
66     ;
67     ;
68     ;
69     ;
70     ;
71     ;
72     ;
73     ;
74     ;
75     ;
76     ;
77     ;
78     ;
79     ;
80     ;
81     ;
82     ;
83     ;
84     ;
85     ;
86     ;
87     ;
88     ;
89     ;
90     ;
91     ;
92     ;
93     ;
94     ;
95     ;
96     ;
97     ;
98     ;
99     ;
100    ;
101    ;
102    ;
103    ;
104    ;
105    ;
106    ;
107    ;
108    ;
109    ;
110    ;
111    ;
112    ;
113    ;
114    ;
115    ;
116    ;
117    ;
118    ;
119    ;
120    ;
121    ;
122    ;
123    ;
124    ;
125    ;
126    ;
127    ;
128    ;
129    ;
130    ;
131    ;
132    ;
133    ;
134    ;
135    ;
136    ;
137    ;
138    ;
139    ;
140    ;
141    ;
142    ;
143    ;
144    ;
145    ;
146    ;
147    ;
148    ;
149    ;
150    ;
151    ;
152    ;
153    ;
154    ;
155    ;
156    ;
157    ;
158    ;
159    ;
160    ;
161    ;
162    ;
163    ;
164    ;
165    ;
166    ;
167    ;
168    ;
169    ;
170    ;
171    ;
172    ;
173    ;
174    ;
175    ;
176    ;
177    ;
178    ;
179    ;
180    ;
181    ;
182    ;
183    ;
184    ;
185    ;
186    ;
187    ;
188    ;
189    ;
190    ;
191    ;
192    ;
193    ;
194    ;
195    ;
196    ;
197    ;
198    ;
199    ;
200    ;
201    ;
202    ;
203    ;
204    ;
205    ;
206    ;
207    ;
208    ;
209    ;
210    ;
211    ;
212    ;
213    ;
214    ;
215    ;
216    ;
217    ;
218    ;
219    ;
220    ;
221    ;
222    ;
223    ;
224    ;
225    ;
226    ;
227    ;
228    ;
229    ;
230    ;
231    ;
232    ;
233    ;
234    ;
235    ;
236    ;
237    ;
238    ;
239    ;
240    ;
241    ;
242    ;
243    ;
244    ;
245    ;
246    ;
247    ;
248    ;
249    ;
250    ;
251    ;
252    ;
253    ;
254    ;
255    ;
256    ;
257    ;
258    ;
259    ;
260    ;
261    ;
262    ;
263    ;
264    ;
265    ;
266    ;
267    ;
268    ;
269    ;
270    ;
271    ;
272    ;
273    ;
274    ;
275    ;
276    ;
277    ;
278    ;
279    ;
280    ;
281    ;
282    ;
283    ;
284    ;
285    ;
286    ;
287    ;
288    ;
289    ;
290    ;
291    ;
292    ;
293    ;
294    ;
295    ;
296    ;
297    ;
298    ;
299    ;
300    ;
301    ;
302    ;
303    ;
304    ;
305    ;
306    ;
307    ;
308    ;
309    ;
310    ;
311    ;
312    ;
313    ;
314    ;
315    ;
316    ;
317    ;
318    ;
319    ;
320    ;
321    ;
322    ;
323    ;
324    ;
325    ;
326    ;
327    ;
328    ;
329    ;
330    ;
331    ;
332    ;
333    ;
334    ;
335    ;
336    ;
337    ;
338    ;
339    ;
340    ;
341    ;
342    ;
343    ;
344    ;
345    ;
346    ;
347    ;
348    ;
349    ;
350    ;
351    ;
352    ;
353    ;
354    ;
355    ;
356    ;
357    ;
358    ;
359    ;
360    ;
361    ;
362    ;
363    ;
364    ;
365    ;
366    ;
367    ;
368    ;
369    ;
370    ;
371    ;
372    ;
373    ;
374    ;
375    ;
376    ;
377    ;
378    ;
379    ;
380    ;
381    ;
382    ;
383    ;
384    ;
385    ;
386    ;
387    ;
388    ;
389    ;
390    ;
391    ;
392    ;
393    ;
394    ;
395    ;
396    ;
397    ;
398    ;
399    ;
400    ;
401    ;
402    ;
403    ;
404    ;
405    ;
406    ;
407    ;
408    ;
409    ;
410    ;
411    ;
412    ;
413    ;
414    ;
415    ;
416    ;
417    ;
418    ;
419    ;
420    ;
421    ;
422    ;
423    ;
424    ;
425    ;
426    ;
427    ;
428    ;
429    ;
430    ;
431    ;
432    ;
433    ;
434    ;
435    ;
436    ;
437    ;
438    ;
439    ;
440    ;
441    ;
442    ;
443    ;
444    ;
445    ;
446    ;
447    ;
448    ;
449    ;
450    ;
451    ;
452    ;
453    ;
454    ;
455    ;
456    ;
457    ;
458    ;
459    ;
460    ;
461    ;
462    ;
463    ;
464    ;
465    ;
466    ;
467    ;
468    ;
469    ;
470    ;
471    ;
472    ;
473    ;
474    ;
475    ;
476    ;
477    ;
478    ;
479    ;
480    ;
481    ;
482    ;
483    ;
484    ;
485    ;
486    ;
487    ;
488    ;
489    ;
490    ;
491    ;
492    ;
493    ;
494    ;
495    ;
496    ;
497    ;
498    ;
499    ;
500    ;
501    ;
502    ;
503    ;
504    ;
505    ;
506    ;
507    ;
508    ;
509    ;
510    ;
511    ;
512    ;
513    ;
514    ;
515    ;
516    ;
517    ;
518    ;
519    ;
520    ;
521    ;
522    ;
523    ;
524    ;
525    ;
526    ;
527    ;
528    ;
529    ;
530    ;
531    ;
532    ;
533    ;
534    ;
535    ;
536    ;
537    ;
538    ;
539    ;
540    ;
541    ;
542    ;
543    ;
544    ;
545    ;
546    ;
547    ;
548    ;
549    ;
550    ;
551    ;
552    ;
553    ;
554    ;
555    ;
556    ;
557    ;
558    ;
559    ;
560    ;
561    ;
562    ;
563    ;
564    ;
565    ;
566    ;
567    ;
568    ;
569    ;
570    ;
571    ;
572    ;
573    ;
574    ;
575    ;
576    ;
577    ;
578    ;
579    ;
580    ;
581    ;
582    ;
583    ;
584    ;
585    ;
586    ;
587    ;
588    ;
589    ;
590    ;
591    ;
592    ;
593    ;
594    ;
595    ;
596    ;
597    ;
598    ;
599    ;
600    ;
601    ;
602    ;
603    ;
604    ;
605    ;
606    ;
607    ;
608    ;
609    ;
610    ;
611    ;
612    ;
613    ;
614    ;
615    ;
616    ;
617    ;
618    ;
619    ;
620    ;
621    ;
622    ;
623    ;
624    ;
625    ;
626    ;
627    ;
628    ;
629    ;
630    ;
631    ;
632    ;
633    ;
634    ;
635    ;
636    ;
637    ;
638    ;
639    ;
640    ;
641    ;
642    ;
643    ;
644    ;
645    ;
646    ;
647    ;
648    ;
649    ;
650    ;
651    ;
652    ;
653    ;
654    ;
655    ;
656    ;
657    ;
658    ;
659    ;
660    ;
661    ;
662    ;
663    ;
664    ;
665    ;
666    ;
667    ;
668    ;
669    ;
670    ;
671    ;
672    ;
673    ;
674    ;
675    ;
676    ;
677    ;
678    ;
679    ;
680    ;
681    ;
682    ;
683    ;
684    ;
685    ;
686    ;
687    ;
688    ;
689    ;
690    ;
691    ;
692    ;
693    ;
694    ;
695    ;
696    ;
697    ;
698    ;
699    ;
700    ;
7
```

```

; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:

```

```

1  APPLICATION NUMBER:  US/09/123,615
2  FILING DATE:
3  PRIOR APPLICATION DATA:
4  APPLICATION NUMBER:  08/705,868
5  FILING DATE:
6  ATTORNEY/AGENT INFORMATION:
7  NAME:  Billings, Lucy J.
8  REGISTRATION NUMBER:  36,749
9  REFERENCE/DOCKET NUMBER:  PF-0117  US
10 TELECOMMUNICATION INFORMATION:
11 TELEPHONE:  415-855-0555
12 TELEFAX:  415-845-4166
13 INFORMATION FOR SEQ ID NO:  4:
14     SEQUENCE CHARACTERISTICS:
15     LENGTH:  388 amino acids
16     TYPE:  amino acid
17     STRANDEDNESS:  single
18     TOPOLOGY:  linear
19     MOLECULE TYPE:  peptide
20     IMMEDIATE SOURCE:
21     LIBRARY:  GenBank
22     CLONE:  1184699
23
24 US-09-123-615-4

```

Query Match	5.9%	Score	132.5	DB	3	Length	388
Best Local	20.8%	Pred. No.	5.2e-06				
Matches	80	Conservative	58	Mismatches	12	Indels	125
						Gaps	20

QY 114 GRRPS-SEAMVGHILRFITFKLQDVENVPLVIQMTDDE--KYLKRLTLDDQ--AYGD 167
 Db 2 GDAPSPKKLH-----LITRNLOEVNGEKKELIKERELKLYGWATTTGKHHVAAYV 54

QY 168 AVENAKDIIACGPDINKTFIFSL-----DYGMSGGYKQNVKIQKH--VTF 213

Db 55 PMSKIAFLKAGEV--TILFALHAYLDNMKAPWELLETLRVSYRYENIKAMTESIGVPL 112

QY 214 NOVKGIFG-----FTDSD-----CIGKISFPALQAA 239
 Db 113 EKFKIKETDYOLSKETVLDVYRLSSVVOHQDSKKAGAEVWQVEHPLRLSGLLPGLOA 171

Qy 240 PFSFNSFPOIFRDRDTIOCLIPCAIDODPYFRMTRDVAPRIGYKPKALHSTFPALOGA 299

Db 172 --LDEEYKLV-----DAQF--GGIDQDKRIFTFAEKYLPALGSKRVHLMNPMPGLTG- 220

0y	300	QTKMSASDPNSSIFLTDTAQIKTKVKNIAF-----	SGGRDTLIEH-----	340
Db	221	-SKMSSSEESKRDLLDKEDYKKLTK-AFCEPQNVENGVSIFIKHLPFLKSEFVL		278

```

OY      341 --RQFGNCDDVDSFMVYLPFLFEDDDKLEQIRKDYTSGAMLTGELKCALIEVLQPLIAEH 398
          : : : : :
DB      279 RDEKMGSG-----KTYTAYV-----LEQDFAAEVHVHPPGQIKNSVEVALNKKL----- 321
          : : : : :

```

```

QY      399 QARKEVTDEIVKEFMTD--RKLS 420
          |  ::|||
DB      322 -----DPIREKFTPALKKLA 337
          |  ::|||

```

RESULT 9
US-08-855-910-11

Sequence 11, Application US/08855910
Patent No. 6221640
GENERAL INFORMATION:
APPLICANT: Tao, Tianshi

APPLICANT: Sassanfar, Mandana
APPLICANT: Gallant, Paul L.
APPLICANT: Shen, Xiaoyu
APPLICANT: through author S

APPLICANT:	Yu, Russell V.
APPLICANT:	Naif, Shamila
TITLE OF INVENTION:	ENTEROCOCCAL AMINOACYL-TRNA SYNTHETASE
TITLE OF INVENTION:	PROBING NUCLEIC ACIDS AND SEQUENCING COMPLEXING SAME

```
;
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
```


ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 STREET: Two Miltia Drive
 CITY: Lexington
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02173
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/855,910
 FILING DATE: 14-MAY-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Brook, David E.
 REGISTRATION NUMBER: 22,592
 REFERENCE/DOCKET NUMBER: CPI95-08
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781) 861-6240
 TELEFAX: (781) 861-9540
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 418 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-855-910-11

Query Match 5.1%; Score 113.5; DB 4; Length 418;
 Best Local Similarity 20.5%; Pred. No. 0.00079;
 Matches 82; Conservative 47; Mismatches 112; Indels 159; Gaps 18;
 111 LTYGRGSSAMAVGHILPFTKMLQDVFNPLVLT----- 146
 33 LYGVDPDGTSMHIGHLIPFMMKRFQLAGHPIYLIGGGGTIGDPSGRTTREVLTQME 92
 147 -----OWTDEKYLW-KDLTLDQAVGAVEN-----AKQIIA 177
 93 AVQHNVDSLNOHKKLFGKAAEVMTVMNNYDMLSELSDLPLRDYGNFNNVMTLAKDIIA 152
 178 CGEDINKFTFSDLDYMGMS-GEYKNNVKIQKVTENVKG----- 218
 153 S--RLBSGISFTERTYQILQSIDFY--TLHKKNIIQLQGGADQMGNTIAGDLIRKKE 207
 219 -----TGF-----DSDICIKISFPALQAFSENSFP--QIFPRDRDIOCLIFCAI 264
 208 GPEAKYFGLTIPMLKADGTRFGKTAGALWLDPKKTSPEFYQFWMLOD----- 258
 265 DQDPYFMTDVAIRIGYPRPALHSTFPALQAGOTKMSASDPNSSIFLTPRAKOIKTK 324
 259 -----RDV--IKYLK-----FTFPLDKER-----IDALAEKYEKE 286
 325 VNRKAFSGSDTIEHRQFGNCDVVSFWYLTFELEDKLEQIRKDYTSGAMLTGELK 384
 287 PKR--EAQRLEAE-----VTRFVHDDALEBAQK--ISEALFSGNIK 326
 385 KALIEVLQ-----PLAEHQARKEVTDEIVKEFMTPRK 418
 327 DLTIETIEQLEHVFV-EITTKAKNIVDWLVDTIEPSK.365

RESULT 10
 US-09-352-990-28
 Sequence 28, Application US/09352990
 Patent No. 6255090
 GENERAL INFORMATION:
 APPLICANT: Farnodu, Layo O.
 APPLICANT: Orozco, Buddy
 APPLICANT: Rafaleki, Antoni
 TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
 ATTORNEY/AGENT INFORMATION:
 FILE REFERENCE: BB-1191

CURRENT APPLICATION NUMBER: US/09/352,990
 CURRENT FILING DATE: 1999-07-14
 EARLIER APPLICATION NUMBER: 60/092,866
 EARLIER FILING DATE: July 15, 1998
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 28
 LENGTH: 377
 TYPE: PRT
 ORGANISM: Synechocystis sp.
 US-09-352-990-28

Query Match 4.9%; Score 109; DB 4; Length 377;
 Best Local Similarity 21.0%; Pred. No. 0.0021;
 Matches 82; Conservative 49; Mismatches 116; Indels 144; Gaps 19;
 91 FSHRDMNQVLDAVENKKPFLVLTGRGSSAMAVGHILPFTKMLQDVFNPLVLTQMTD 150
 33 FAHRSTTAM-----DKPRIL-SGVQPTGN-LHLGNVLGAI-RSWEQ-----QOHY 75
 151 DEKYLKDL-----TLDOAVGDAVENAKDIIAGFDPINKTFIF----- 188
 76 DNFECVVDLHAITVPHNPOTLQO--DTITIALYIACSIDIQSTIFVQSHVAHSELA 132
 189 -----SDLDYMGSSGFYKNNVKIQKVTENVKIGFPTSDICIKISFPALQAPSF 242
 133 WLNCVTPLNWLBERMIOFKEKAVKQGENVS-----VGLLDYFVLMAA--- 174
 243 SNSFPQIFPRDRDIOCL-----IPCAIDDPYFRMTRDV-----ARIGYPKRA 286
 175 -----DILLYDAKVPVGEQKQHELTRDIYIRINDKRGREDATVLTLPBL 222
 287 L-LHSTFPALQAGOTKMSASDPN--SIFLTPRAKOIKTKYKNAFSGGRDTIEHRQF 343
 223 IRREGARVNSLADGTYKMSKSDSELSRLNLDPRPMIKKVKK----- 266
 344 GNCNDVVSFWYLTFELEDKLEQIRKDYTSGAMLTGELKAL----- 387
 267 ---CKTDPO-RGLWF---DDPERBECHNLLTYLLTSLNQTKEAVVQBCAEMGNGQPKPL 319
 388 -----IEVLOPLAEHQARKEVTDEIVK 411
 320 TETAIALBPQAKYAEIILADRGL-DRILIQ 349

RESULT 11
 US-08-923-867-2
 Sequence 2, Application US/08923867
 Patent No. 5851809
 GENERAL INFORMATION:
 APPLICANT: Lawlor, Elizabeth
 TITLE OF INVENTION: NOVEL TRYPTOPHANYL tRNA
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham Corporation
 STREET: 709 Swedeland Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19406-0939
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/923,867
 FILING DATE: 04-SEP-1997
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 9619072.3
 FILING DATE: 12-SEP-1996
 ATTORNEY/AGENT INFORMATION:

NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31624
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-923-867-2

Query Match
Best Local Similarity 25.3%; Score 108; DB 2; Length 197;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 260 IPCAIDDDPYFRTRD-----VAPRIGYPPKALLHSTFPALQGAQTKMSA 305
DB 5 VEVGTDOQKMIETREIIVSFNNAYNCVIVPEEGIYENE--RAGRLEGLDG-NAKMSK 61

QY 306 SDPNSSIFLDTAKQITKYNKHAFGSGRDTIEHRQFGNCDVVSFWYLTFF--LEDD 363
DB 62 S-LNNGIYIADADLTARKKVMSTYDPDHIRVEDPKIEGN---MVFFYLDVFGRPEDA 116

QY 364 DKLEQIRKDYTGAMLTGELKALILEVLOPLIAEHQARKEVTDEI 409
DB 117 QETADMKERYQRGGLGDVYTKRYLLEILEREIGPIRERIRIEPAKDM 162

RESULT 12
US-08-928-100-4
Sequence 4, Application US/08928100
Patent No. 6046174
GENERAL INFORMATION:
APPLICANT: Gentry, Danlie
APPLICANT: Greenwood, Claire
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: NO. 6046174el trps
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,100
FILING DATE: 12-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9619072.3
FILING DATE: 12-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31624-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-928-100-4

Query Match
Best Local Similarity 25.3%; Score 108; DB 3; Length 197;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 260 IPCAIDDDPYFRTRD-----VAPRIGYPPKALLHSTFPALQGAQTKMSA 305
DB 5 VEVGTDOQKMIETREIIVSFNNAYNCVIVPEEGIYENE--RAGRLEGLDG-NAKMSK 61

QY 306 SDPNSSIFLDTAKQITKYNKHAFGSGRDTIEHRQFGNCDVVSFWYLTFF--LEDD 363
DB 62 S-LNNGIYIADADLTARKKVMSTYDPDHIRVEDPKIEGN---MVFFYLDVFGRPEDA 116

QY 364 DKLEQIRKDYTGAMLTGELKALILEVLOPLIAEHQARKEVTDEI 409
DB 117 QETADMKERYQRGGLGDVYTKRYLLEILEREIGPIRERIRIEPAKDM 162

RESULT 13
US-09-183-134-2
Sequence 2, Application US/09183134
Patent No. 615759
GENERAL INFORMATION:
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: NOVEL TRYPTOPHANYL CRNA
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/183,134
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/923,867
FILING DATE: 04-SEP-1997
APPLICATION NUMBER: 9619072.3
FILING DATE: 12-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31624
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-183-134-2

Query Match
Best Local Similarity 25.3%; Score 108; DB 4; Length 197;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

```

QY 260 PCALIDODDYFMTSD-----VPRIGIPKPAHSHFFPPLQAOAQIKMSA 305
Db 5 VPGVTDQKPMITQTRIEIVS FNNANVCDLVVEBEGITPENE--RAQRLPELSD--NAIMSK 61
QY 306 SDPNSSIFLTPAKOIKTKYVNHKHAFFSGGRDTEIEHRQFGNCVDVSPFWLTFE--LEDD 363
Db 62 S-LINGIYIADADYILRKVKVMSMYTDPDRIYVDDPGKIEGN----MVFRILVDFGRPEDA 116
QY 364 DKLEOIRKDYDTSGAMLTGELKALIEVLPLAEHQARRKEVDEI 409
Db 117 QEIDAMKERYQRGGLGDVTKRYLLEILRELBQIPNERRIEFAKDM 162

```

RESULT 14
US-09-492-581-4
; Sequence 4, Application US/09492581

APPLICANT: Gentry, Danile
 APPLICANT: Greenwood, Claire
 APPLICANT: Lawlor, Elizabeth
 TITLE OF INVENTION: No. 6346409el trps
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham Corporation
 STREET: 709 Swedeland Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19406-0939
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/492,581
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,100
 FILING DATE: 12-SEP-1997
 APPLICATION NUMBER: 9619072.3
 FILING DATE: 12-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Gimm, Edward R
 REGISTRATION NUMBER: 38,891
 REFERENCE/DOCKET NUMBER: P31624-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090
 TELEX:
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 197 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-492-581-4

	Query Match	4.8%	Score 108,	DB 4;	Length 197,	:
	Best Local Similarity	25.3%;	Pred. No. 0.00092;			
	Matches	42;	Mismatches	73;	Indels	24; Gaps
						6.
Qy	260	IPCAIDDDPYFRMTD-----VAPRIGYPRKALHSTFPFPALOGAQTMSA	305			
	:	:	:	:	:	:
Db	5	VPGVTDOQPMIEQTREIVRSFNNAACDVLEPEGIYPENE--RAGRIGLGDG-NARKSK	61			
	:	:	:	:	:	:
Qy	306	SDPNSIFLTTPAKQIKTKVNKHAFSGGRDTIEEHRQFGNCDVDVSFEWLYTFE--LEED	363			
	:	:	:	:	:	:
Db	62	S-LNNGIYLAADADQLTRKKVMSSMYTDPHIRIREDQCKLEGN---MVFFYLDFVRPEDA	116			

```

QY      364 DKLEQIRKDYTSGAMLTGELKKALIEVLQPIAHHQARRKEVTDEI 4093
      :: :: | | : : | : | : : : | | : :
Db     117 QEIADMKERYQRGGLGVYTKRYLLEILERELGPIRERRIEFAKDM 1622

```

RESULT 15
US-09-425-666-4
; Sequence 4, Application US/09425666
; Patent No. 6416976
; Patent No. 6416976

```

1 GENERAL INFORMATION:
2 APPLICANT: Gentry, Danile
3 APPLICANT: Greenwood, Claire
4 APPLICANT: Lawlor, Elizabeth
5 TITLE OF INVENTION: NO. 6416976el tips
6 NUMBER OF SEQUENCES: 6
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: SmithKline Beecham Corporation
9 STREET: 709 Swedeland Road
10 CITY: King of Prussia
11 STATE: PA
12 COUNTRY: USA
13 ZIP: 19406-0939
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette
16 COMPUTER: IBM Compatible
17 OPERATING SYSTEM: DOS
18 SOFTWARE: FASTSEQ for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/09/425,666
21 FILING DATE:
22 CLASSIFICATION:
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: 08/928,100
25 FILING DATE:
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Glimm, Edward R
28 REGISTRATION NUMBER: 38,891
29 REFERENCE/DOCKET NUMBER: P31624-1
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 610-270-4478
32 TELEFAX: 610-270-5090
33 TELEX:
34 INFORMATION FOR SEQ ID NO: 4:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 197 amino acids
37 TYPE: amino acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: protein
41 US-09-425-666-4

```

Query Match	4.8%;	Score 108;	DB 4;	Length 197;
Best Local Similarity	25.3%;	Pred. No. 0.00092;		
Matches	42;	Conservative	27;	Mismatches 73; Indels 24; Gaps 6

QY	260	IPCAIDDDPYFRMTRD-----	-VAPRIGYKPALHSTFPALQGAQTKMSA	305
Db	5	VPVGTDDKPMLEGRREIYRSFNANVNCVLYVEPSGIVENE--	RAGRLPGLDG-VAKKSK	61
QY	306	SDPNSSTFLDTAKQIKTKONKHAASGGGRDTIEEHROFGGNCDDVVSMTLTF--	LEDD	363
Db	62	S-LNNGIYILDADDAOTLRKKVMSMYTDPDPIHIVEDPGKIEGN---	MVFHYLDVFGRPEDA	116
QY	364	DKLEQIRKDYTSGAMLTGELKKALIEVYQPLAEHQARKVETDEI	409	
Db	117	QEIDMKERYQGGELGVKTKRYLLEIERELGPIRRIERIFAKDM	162	

Search completed: July 10, 2003, 12:32:50
Job time : 15.8485 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:28:00 ; Search time 16.4461 Seconds
(without alignments)
2478.466 Million cell updates/sec

Title: US-09-813-718-10_COPY_48_471

Perfect score: 2246
Sequence: 1 MSYKRAAGEDYKADCEPQNP.....VTDEIVKEMTPRKLSFDFQ 424

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:*

1: p1r1:.*
2: p1r2:.*
3: p1r3:.*
4: p1r4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2231	99.3	471	1 A41706	tryptophan-tRNA 11
2	2134.5	95.0	475	1 YWBO	tryptophan-tRNA 11
3	2041	90.9	475	1 YMRBPR	tryptophan-tRNA 11
4	2032	90.5	481	2 S50053	tryptophan-tRNA 11
5	1210	53.9	395	2 S58157	hypothetical prote
6	1163	51.8	432	2 S51901	tryptophan-tRNA 11
7	907	40.4	386	2 C90190	tryptophan-tRNA
8	803	35.8	385	2 C75020	tryptophanyl-tRNA
9	626.5	27.9	301	2 G71206	tryptophan-tRNA 11
10	534.5	23.8	380	2 G84373	tryptophanyl-tRNA
11	409.5	18.2	370	2 F64476	tryptophan-tRNA 11
12	397.5	17.7	364	2 E69131	tryptophan-tRNA 11
13	386	17.2	134	2 T43806	tryptophan-tRNA 11
14	370.5	16.5	420	2 E69461	tryptophanyl-tRNA
15	353	15.7	374	2 D72477	probable tryptopha
16	274.5	12.2	513	2 F84371	tryptophan-tRNA
17	192	8.5	323	2 H69346	tyrosyl-tRNA synth
18	188	8.4	364	2 E72512	probable tyrosyl-t
19	185.5	8.3	341	2 D95260	tryptophanyl-tRNA
20	185.5	8.3	341	2 G98125	tryptophan-tRNA 11
21	174.5	7.8	341	2 B86633	tryptophan-tRNA 11
22	172	7.7	394	2 A45999	tyrosine-tRNA 11ga
23	170.5	7.6	366	2 S75410	tyrosine-tRNA 11ga
24	169	7.5	395	2 H70385	tryptophan-tRNA 11
25	167	7.4	351	2 E75438	tryptophanyl-tRNA
26	166.5	7.4	346	2 B71496	tryptophan-tRNA 11
27	164.5	7.3	895	2 A86410	protein F3M18.22 1
28	163.5	7.3	353	2 E70100	tryptophan-tRNA 11
29	163	7.3	460	2 C84750	probable tyrosyl-t

30	159.5	7.1	337	2 F71300	tryptophan-tRNA 11
31	158.5	7.1	375	2 B75072	tyrosyl-tRNA synth
32	158	7.0	345	2 C91654	tryptophanyl-tRNA
33	154.5	6.9	102	2 T44994	probable tryptopha
34	152	6.8	344	2 H86590	tryptophanyl tRNA
35	152	6.8	344	2 C72034	tryptophan-tRNA 11
36	149.5	6.7	327	2 C84374	tyrosyl-tRNA synth
37	149.5	6.7	408	2 T03741	probable tyrosine-
38	146	6.5	339	2 E64676	tryptophan-tRNA 11
39	146	6.5	375	2 F71093	tyrosine-tRNA 11ga
40	145	6.5	337	2 A11066	tryptophan-tRNA 11
41	138	6.1	319	2 H69102	tyrosine-tRNA 11ga
42	135.5	6.0	328	2 C72370	tryptophan-tRNA 11
43	135	6.0	365	2 E82052	tryptophanyl-tRNA
44	132.5	5.9	343	2 S73024	tryptophan-tRNA 11
45	132	5.9	328	1 YWBSF	tryptophan-tRNA 11

ALIGNMENTS

RESULT 1

A41706
tryptophan-tRNA ligase (EC 6.1.1.2) [similarity] - human
N:Alternate names: interferon-inducible protein IFP53; peptide-chain release factor homo)
C:Species: Homo sapiens (man)
C>Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 03-Jun-2002
C/Accession: A41633; A41706; S19246; J06076; J0533; S26287
R:Elekter, J.; Rasmussen, H.H.; Jussensen, J.
Proc. Natl. Acad. Sci. U.S.A. 88, 11520-11524, 1991
A>Title: Human interferon gamma potentially induces the synthesis of a 55-kDa protein (gamma
A:Reference number: A41633; MUID:92107982; PMID:1763065
A:Accession: A41633
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-471 <RUB>
A:Cross-references: GB:X59892; NID:930820; PIDN:CAA42545.1; PID:930821
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.
J. Biol. Chem. 266, 24245-24248, 1991
A>Title: Interferon induces tryptophanyl-tRNA synthetase expression in human fibroblasts.
A:Reference number: A41706; MUID:92105071; PMID:1761529
A:Accession: A41706
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-471 <RUB>
A:Cross-references: GB:M77804; NID:9184656; PIDN:AAA67324.1; PID:9184657
R:Buwilt, U.; Flohr, T.; Boettger, E.C.
EMBO J. 11, 489-496, 1992
A>Title: Molecular cloning and characterization of an interferon induced human cDNA with
A:Reference number: S19246; MUID:92164636; PMID:1537332
A:Accession: S19246
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-423, 'R', 425-471 <BUW>
A:Cross-references: EMBL:X62570; NID:932708; PIDN:CAA44450.1; PID:932709
A>Note: 213-ser and 214-tyr were also found
R:Prolova, L.Y.; Grigorova, A.Y.; Sudomolina, M.A.; Kiselev, L.L.
Gene 128, 237-245, 1993
A>Title: The human gene encoding tryptophanyl-tRNA synthetase: Interferon-response element
A:Reference number: J06076; MUID:93292992; PMID:7685728
A:Accession: J06076
A:Molecule type: DNA
A:Residues: 1-111,182-471 <PROL>
A:Cross-references: GB:X67918; GB:S62837; NID:937968; GB:X67919; NID:937969; GB:X67920; 1
4; NID:937974; GB:X67925; GB:S62855; NID:937975; GB:X67926; GB:S62856; NID:937976; GB:X6
A>Note: the authors translated the codon GGG for residue 55 as Cys and GAG for residue 34
A>Note: this translation is not annotated in Genbank entries HSMRX11, HSMRX18, HSMRX2,
R:Prolova, L.Y.; Sudomolina, M.A.; Grigorova, A.Y.; Zinovleva, O.L.; Kiselev, L.L.
Gene 109, 291-296, 1991
A>Title: Cloning and nucleotide sequence of the structural gene encoding for human trypt
A:Reference number: J0533; MUID:92112058; PMID:1765274
A:Accession: J0533
A:Molecule type: mRNA

A:Residues: 1-212; 'GD', 215-471 <PRO2>
 A/Cross-references: GB:M61715; NID:g340367; PIDN:AAA61298.1; PID:g340368
 A/Experimental source: /libroblast
 C/Genetics:
 A/Gene: GDB:WARS; IFF53
 A/Cross-references: GDB:119632; OMIM:191050
 A/Map position: 14q23-14q31
 A/Intons: 33/3; 105/1; 141/2; 181/2; 242/2; 276/1; 313/3; 371/3; 418/3
 C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
 C/Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis
 F:19-64/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 99.3%; Score 2231; DB 1; Length 471;
 Best Local Similarity 99.5%; Pred. No. 1.8e-173; Indels 0; Gaps 0;

Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSYKAAAGEDYKADCPGPNPAPTSNHPDTEAEDEFPDPTVQTSASAKIDYKLIYRF 60
 DB 48 MSYKAAAGEDYKADCPGPNPAPTSNHPDTEAEDEFPDPTVQTSASAKIDYKLIYRF 107
 QY 61 GSKIDELINIRERATGQRPVPLVIRGIFPSHRDNNQVLDAYENKKPFYLYTGRGSSSE 120
 DB 108 GSKIDELINIRERATGQRPVPLVIRGIFPSHRDNNQVLDAYENKKPFYLYTGRGSSSE 167
 QY 121 AMVGHILIPFTKMLQDVFNPLVIRGIFPSHRDNNQVLDAYENKKPFYLYTGRGSSSE 180
 DB 168 AMVGHILIPFTKMLQDVFNPLVIRGIFPSHRDNNQVLDAYENKKPFYLYTGRGSSSE 227
 QY 181 DINKTFISDLIDYMGSSGFYKVVKIQKVTENQVKGIFGFTSDCIGKISFPAIOAAS 240
 DB 228 DINKTFISDLIDYMGSSGFYKVVKIQKVTENQVKGIFGFTSDCIGKISFPAIOAAS 287
 QY 241 SFNSFPQIFRDRDIOCLIPCAIDQDPYFRMTDVAIRGYPFPALHSTFFPALQGAQ 300
 DB 288 SFNSFPQIFRDRDIOCLIPCAIDQDPYFRMTDVAIRGYPFPALHSTFFPALQGAQ 347
 QY 301 TKMSASDPSNSIFLITDPAKQIKTKVNGHAFSGGSDTIEHRQFGNCDDVVSFMYLTFEFL 360
 DB 348 TKMSASDPSNSIFLITDPAKQIKTKVNGHAFSGGSDTIEHRQFGNCDDVVSFMYLTFEFL 407
 QY 361 EDDDKLEQIRKDYTSGLMTEGELKALIEVLQPIAIAHQARKKEVTDEIYKEFMTPKLS 420
 DB 408 EDDDKLEQIRKDYTSGLMTEGELKALIEVLQPIAIAHQARKKEVTDEIYKEFMTPKLS 467
 QY 421 FDFQ 424
 DB 468 FDFQ 471

RESULT 2

YMRBO
 tryptophan-tRNA ligase (EC 6.1.1.2) [validated] - bovine

N/Alternate names: tryptophanyl-tRNA synthetase
 C/Species: Bos primigenius taurus (cattle)
 C/Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 03-Jun-2002

A/Accession: A40279; JN0354; S10450; S14540
 A/Reference number: A40379; MUID:9132348; PMID:1907847

A/Title: A mammalian tryptophanyl-tRNA synthetase shows little homology to prokaryotic
 A/Accession: A40279
 A/Residues: 1-475 <GNR>

A/Cross-references: GB:M74074; EMBL:X53918; NID:g163798; PIDN:AAA30799.1; PID:g163799
 A/Experimental source: pancreas

A/Note: the authors translated the codon CTG for residue 347 as Ala and CAG for residue
 A/Note: part of this sequence was confirmed by protein sequencing
 R:Zarganova, T.A.; Kovalova, G.K.; Favorova, O.O.; Aeyina, H.B.; Telezhinskaya, I.H.

A/Title: Amino acid sequence of several peptides of tryptophanyl-tRNA synthetase from ca
 A/Reference number: JN0354; MUID:90211408; PMID:2631664

A/Accession: JN0354
 A/Molecule type: protein

A:Residues: 112-124;282-287, 'N', 288, 'F', 289-292, 'Q', 293-294, 'IR', 336-353; 423-441, 443-449,
 A/Experimental source: liver
 A/Note: this paper is in Russian
 C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
 C/Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis
 F:24-69/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 95.0%; Score 2134.5; DB 1; Length 475;
 Best Local Similarity 95.0%; Pred. No. 1.3e-165; Indels 1; Gaps 1;

Matches 402; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

QY 2 SYKAAAGEDYKADCPGPNPAPTSNHPDTEAEDEFPDPTVQTSASAKIDYKLIYRF 61
 DB 54 SYKAAAGEDYKADCPGPNPAPTSNHPDTEAEDEFPDPTVQTSASAKIDYKLIYRF 113
 QY 62 SSKIDELINIRERATGQRPVPLVIRGIFPSHRDNNQVLDAYENKKPFYLYTGRGSSSE 121
 DB 114 SSKIDELINIRERATGQRPVPLVIRGIFPSHRDNNQVLDAYENKKPFYLYTGRGSSSE 173
 QY 122 MHVGHILIPFTKMLQDVFNPLVIRGIFPSHRDNNQVLDAYENKKPFYLYTGRGSSSE 181
 DB 174 MHVGHILIPFTKMLQDVFNPLVIRGIFPSHRDNNQVLDAYENKKPFYLYTGRGSSSE 232
 QY 182 INKTFISDLIDYMGSSGFYKVVKIQKVTENQVKGIFGFTSDCIGKISFPAIOAAS 241
 DB 233 INKTFISDLIDYMGSSGFYKVVKIQKVTENQVKGIFGFTSDCIGKISFPAIOAAS 292
 QY 242 SFNSFPQIFRDRDIOCLIPCAIDQDPYFRMTDVAIRGYPFPALHSTFFPALQGAQ 301
 DB 293 SFNSFPQIFRDRDIOCLIPCAIDQDPYFRMTDVAIRGYPFPALHSTFFPALQGAQ 352
 QY 302 KMSASDPSNSIFLITDPAKQIKTKVNGHAFSGGSDTIEHRQFGNCDDVVSFMYLTFEFL 361
 DB 353 KMSASDPSNSIFLITDPAKQIKTKVNGHAFSGGSDTIEHRQFGNCDDVVSFMYLTFEFL 412
 QY 362 DDDKLEQIRKDYTSGLMTEGELKALIEVLQPIAIAHQARKKEVTDEIYKEFMTPKLS 421
 DB 413 DDDKLEQIRKDYTSGLMTEGELKALIEVLQPIAIAHQARKKEVTDEIYKEFMTPKLS 472
 QY 422 DDFQ 424
 DB 473 DDFQ 475

RESULT 3

YMRBRP

tryptophan-tRNA ligase (EC 6.1.1.2) [validated] - rabbit

N/Alternate names: tryptophanyl-tRNA synthetase
 C/Species: Oryctolagus cuniculus (domestic rabbit)
 C/Date: 30-Sep-1992 #sequence_revision 13-Feb-1998 #text_change 03-Jun-2002

A/Accession: A35904; S37396
 A/Reference number: A35904; MUID:90239043; PMID:2185472

A/Title: Cloning and expression of a mammalian peptide chain release factor with sequence
 A/Accession: A35904
 A/Molecule type: mRNA

A/Residues: 1-475 <LBR>
 A/Cross-references: GB:M33460
 R:Polova, L.Y.; Dalphin, M.E.; Justesen, J.; Powell, R.J.; Drugeon, G.; McCaughan, K.K.;

EMBO J. 12, 4013-4019, 1993
 A/Title: Mammalian polypeptide chain release factor and tryptophanyl-tRNA synthetase are
 A/Reference number: S37396; MUID:94009008; PMID:8404667

A/Accession: S37396
 A/Molecule type: mRNA

A/Residues: 166-177 <PRO>
 C/Genetics:
 A/Gene: WRS

C/Complex: homodimer [validated, MUID:94009008]
 A/Description: EC 6.1.1.2 [validated, MUID:94009008]; catalyzes the ATP-dependent formati
 A/Note: mammalian WRS (tryptophanyl-tRNA synthetase) and eRF (polypeptide chain release

C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

C:Keywords: aminoacyl-tRNA synthetase; ATP; homodimer; ligase; metalloprotein; protein h
 F:23-68/Domain: amino acid-tRNA ligase repeat homology <ATL>
 F:174-177/Region: ATP-binding motif (HKGH)

Query Match 90.9%; Score 2041; DB 1; Length 475;
 Best Local Similarity 89.8%; Pred. No. 5,1e-158;
 Matches 380; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

2 SYKAAAGEDYKADCPGPNAPTSNHGPDATAEEDPVDPWTQTSSAKGIDYDKLIYRF 61
 |||
 53 SYKAAAGEDYKADCPGPNAPTSNHGPDATAEEDPVDPWTQTSSAKGIDYDKLIYRF 112
 |||
 62 SKKIDKELINIRERATGQRPHFLRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSPA 121
 |||
 113 SKKIDKELINIRERATGQRPHFLRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSPA 172
 |||
 122 MHVGHILPFIPTKMLQDVFNVPVLIQMTDDEKYLMDLTLDQAYGDAVENAKDIIACGF 181
 |||
 173 MHVGHILPFIPTKMLQDVFNVPVLIQMTDDEKYLMDLTLDQAYGDAVENAKDIIACGF 232
 |||
 182 INKTIFFSLDYMGMSSGFYKNNVKIQKHTFNQVKGIFGFTSDCIKISFPALQAAS 241
 |||
 223 VNKTIFFSLDYMGMSSGFYKNNVKIQKHTFNQVKGIFGFTSDCIKISFPALQAAS 292
 |||
 242 FSNSEFQIFRDRDIOCLIPCAIDODPYFRMTRDVAPRIGYKPKALHSTFPALQAQ 301
 |||
 293 FSNSEFQIFRDRDIOCLIPCAIDODPYFRMTRDVAPRIGYKPKALHSTFPALQAQ 352
 |||
 302 KMSASDPNSSIFLTDPAKQIKTKVKNKAFSGGRDTIEHRQFGNCVDVDSFMYLTFFLE 361
 |||
 353 KMSASDPNSSIFLTDPAKQIKTKVKNKAFSGGRDTIEHRQFGNCVDVDSFMYLTFFLE 412
 |||
 362 DDDKLEQIRKDYTSGLMTELKKAILEVLOPLIAEHQARKKEVDEIYKEMTPRKLSF 421
 |||
 413 DDDKLEQIRKDYTSGLMTELKKAILEVLOPLIAEHQARKKEVDEIYKEMTPRKLSF 472
 |||
 422 DFG 424
 |||
 473 HQ 475

RESULT 4

S50053
 C:Species: Mus musculus (house mouse)
 C:Date: 07-May-1995 #sequence_revision 10-Nov-1995 #text_change 03-Jun-2002
 C:Accession: S50053; S50052; I49391; S31461; S31462
 R:Pajot, B.; Sarger, C.; Bonnet, J.; Garret, M.
 J. Mol. Biol. 242, 599-603, 1994
 A>Title: An alternative splicing modifies the C-terminal end of tryptophanyl-tRNA synthe
 A:Reference number: S50052; MID:95018226; PMID:7932716
 A:Accession: S50053
 A:Molecule type: mRNA
 A:Residues: 1-481 <PAJ>
 A:Cross-references: EMBL:X69657; NID:g55437; PIDN:CAA49348.1; PID:g55438
 A:Genetics: LSF
 A>Note: Intron position was determined by sequencing of genomic DNA
 A:Accession: S50052
 A:Molecule type: mRNA
 A:Residues: 1-475 <PAM>
 A:Cross-references: EMBL:X69656; NID:g55435; PIDN:CAA49347.1; PID:g55436
 A:Genetics: SSF
 R:Kiselev, L.L.
 Biochimie 75, 1027-1039, 1993
 A>Title: Mammalian tryptophanyl-tRNA synthetases.
 A:Reference number: I49391; MID:94257729; PMID:7515282
 A:Accession: I49391
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-481 <RES>
 A:Cross-references: EMBL:X69657; NID:g55437; PIDN:CAA49348.1; PID:g55438
 C:Genetics: <LSF>
 A:Gene: WRS

A:introns: 475/2
 A>Note: The list of introns may be incomplete; clone W13
 C:Genetics: <SSF>
 A:Gene: WRS

A:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
 C:Keywords: alternative splicing; aminoacyl-tRNA synthetase; ligase; protein biosynthesis
 F:1-481/Product: tryptophan-tRNA ligase alpha-2 chain long splice form #status predicted
 F:1-475/Product: tryptophan-tRNA ligase alpha-2 chain short splice form #status predicted
 F:23-68/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 90.5%; Score 2032; DB 2; Length 481;
 Best Local Similarity 89.6%; Pred. No. 2.8e-157;
 Matches 380; Conservative 22; Mismatches 22; Indels 0; Gaps 0;

1 MSYKAAAGEDYKADCPGPNAPTSNHGPDATAEEDPVDPWTQTSSAKGIDYDKLIYRF 60
 |||
 52 MSYKAAAGEDYKADCPGPNAPTSNHGPDATAEEDPVDPWTQTSSAKGIDYDKLIYRF 111
 |||
 61 GSKKIDKELINIRERATGQRPHFLRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSPA 120
 |||
 112 GSKKIDKELINIRERATGQRPHFLRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSPA 171
 |||
 121 AMHVGHLIPFIPTKMLQDVFNVPVLIQMTDDEKYLMDLTLDQAYGDAVENAKDIIACGF 180
 |||
 172 AMHVGHLIPFIPTKMLQDVFNVPVLIQMTDDEKYLMDLTLDQAYGDAVENAKDIIACGF 231
 |||
 181 DINKTIFFSLDYMGMSSGFYKNNVKIQKHTFNQVKGIFGFTSDCIKISFPALQAAP 240
 |||
 222 DINKTIFFSLDYMGMSSGFYKNNVKIQKHTFNQVKGIFGFTSDCIKISFPALQAAP 291
 |||
 241 SFSNPFQIFRDRDIOCLIPCAIDODPYFRMTRDVAPRIGYKPKALHSTFPALQAQ 300
 |||
 292 SFSNPFQIFRDRDIOCLIPCAIDODPYFRMTRDVAPRIGYKPKALHSTFPALQAQ 351
 |||
 301 TKMSASDPNSSIFLTDPAKQIKTKVKNKAFSGGRDTIEHRQFGNCVDVDSFMYLTFFLE 360
 |||
 352 TKMSASDPNSSIFLTDPAKQIKTKVKNKAFSGGRDTIEHRQFGNCVDVDSFMYLTFFLE 411
 |||
 361 EDDKLEQIRKDYTSGLMTELKKAILEVLOPLIAEHQARKKEVDEIYKEMTPRKLS 420
 |||
 412 EDDKLEQIRKDYTSGLMTELKKAILEVLOPLIAEHQARKKEVDEIYKEMTPRKLS 471
 |||
 421 DFG 424
 |||
 472 HQ 475

RESULT 5

S58157
 C:Species: Schizosaccharomyces pombe
 C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 10-Dec-1999
 C:Accession: S58157; T38561
 R:Gentles, S.; Churcher, C.M.
 submitted to the EMBL Data Library, July 1995
 A:Reference number: S58145
 A:Accession: S58157
 A:Molecule type: DNA
 A:Residues: 1-395 <GEN>
 A:Cross-references: EMBL:Z50142; NID:g1052783; PIDN:CAA90500.1; PID:g1052796
 R:Gentles, S.; Churcher, C.M.; Barrett, B.G.; Rajandream, M.A.; Wain, S.V.
 submitted to the EMBL Data Library, July 1995
 A:Accession: T38561
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-395 <GB2>
 A:Cross-references: EMBL:Z50142; PIDN:CAA90500.1; GSPDB:GNO0066; SPDB:SPAC2F7.13C
 A:Experimental source: strain 972h-; cosmid c2f7
 C:Genetics:
 A:Gene: SPDB:SPAC2F7.13C

A:Map position: 1
C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 53.9%; Score 1210; DB 2; Length 395;
Best Local Similarity 59.6%; Pred. No. 1.5e-90;
Matches 235; Conservative 60; Mismatches 91; Indels 8; Gaps 4;

QY 34 EEDFVDPWTVQTS-----SAGIDYDKLIVFGSSKIDKELINRIERATGGRPHHFRGI 89
DB 4 EEOITVPMVKGSTVDEBEGKIDYERLIVQFGTKITPEOLERPEKLTGKKPHLLRGA 63

QY 90 FFSHRDMNOVLDAVENKKPFYLYTGRGSSSAMVGHILPFIPTKMLQDVFNPLVLTOMT 149
DB 64 FFSHRDPMILDRERQKKPFYLYTGRGSSSMHGMIFPMCKMLQDVFNPLVLTQLT 123

QY 150 DDEKYLMD-LTLDOAYGDAVENAKDIIACGFDINKTFISDLDYMGSSGFYKNVVKIQ 208
DB 124 DDEKFLFKQGVSLDDCCRFARENKDIIVAGFDPKTFIFMNSTYVG--GAFYQNVVRIA 181

QY 209 KHYTFNQVKGIFGFTSDICIGISFPALQAPSPNSFPQIFRRTDIOCLIPALIDDP 268
DB 182 KCITANSGKCFGFTSDISIGKIHFAIQAAPSSSPFPHFNGAKDIPCLIPALIDDP 241

QY 269 YFMTBDVAPRIGYPRKALHSTFEPALQAGQTMASDPNSSIFLDTAKQIKTKVKNK 328
DB 242 YFPLTRVSGRLKPKKALHSHRFPALQGPQSGMSKSSAIFMTDTNKKIKNKNRH 301

QY 329 AFSGGRDTEIEHROFGNCVDVSPMYLTFLEDDKLEQIRKDYSGAMLTKELKALI 388
DB 302 AFSGGATIEIHRKGGNPVDVAYQYLSFLDDDEKLEKQLYMYTKAGTLSTGEMKECI 361

QY 389 EVLQPLAEHQARKEVTDEIVKEFMT-PRKLSF 421
DB 362 KLLQGVSDFOAKRSKVDKATLDMFMDSKRLKEM 395

RESULT 6
S51901
tryptophan-tRNA ligase (EC 6.1.1.2) [similarity] - yeast (*Saccharomyces cerevisiae*)
N/Alternate names: protein HR432; protein O0792; tryptophanyl-tRNA synthetase
C/Species: *Saccharomyces cerevisiae*
C/Date: 05-May-1995 #sequence revision 03-Aug-1995 #text_change 03-Jun-2002
A/Accession: S51901; S59177; S66793
R/Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.
submitted to the EMBL Data Library, January 1995
A/Description: Sequence analysis of a 44kb DNA fragment of yeast chromosome XV including and a Delta.
A/Reference number: S51848
A/Accession: S51901
A/Molecule type: DNA
A/Residues: 1-432 <VAM>
A/Cross-references: EMBL:Z48149; NID:g663234; PIDN:CAA88164.1; PID:g663256
R/Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.
Yeast 11, 1069-1075, 1995
A/Title: Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV including the delta element.
A/Reference number: S59156; MUID:96076631; PMID:7502582
A/Accession: S59177
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-432 <VAM>
A/Cross-references: EMBL:Z48149; NID:g663234; PIDN:CAA88164.1; PID:g663256
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
R/Durand, P.; Hilger, F.; Portetelle, D.; Vandenbol, M.
submitted to the Protein Sequence Database, July 1996
A/Reference number: S66791
A/Accession: S66793
A/Molecule type: DNA
A/Residues: 1-432 <DUR>
A/Cross-references: EMBL:Z74839; NID:g1419947; PIDN:CAA99110.1; PID:g1419948; GSPDB:GN00
A/Experimental source: strain S286C
C/Genetics:
A/Gene: SGD:WRS1; WRS1; MIPS:YOL097C

A/Cross-references: SGD:S0005457
A:Map position: 15L
C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
C/Species: *Sulfolobus solfataricus*
C/Keywords: ligase
F/117-120/Region: ATP-binding motif (HXGH)

Query Match 51.8%; Score 1163; DB 2; Length 432;
Best Local Similarity 54.8%; Pred. No. 1.1e-86;
Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

QY 30 ATEAEDFVDPWTV-----QTSAGIDYDKLIVFGSSKIDKELINRIERATGGRPHHFR 84
DB 19 STDVKEQVTVPMVEGSGVDEGGRANIDYDKLIDKQGTGRVNEETLKRKQVTRGRRPHF 78

QY 85 LRGRIFPSHRDMNOVLDAVENKKPFYLYTGRGSSSAMVGHILPFIPTKMLQDVFNPLV 144
DB 79 LRGLFSEEDFKLIDLEQGFPLLYTGRGSSSMHGMIPVFTKMLQDVDPVL 138

QY 145 VIQMTDDEKYLMD-LTLDOAYGDAVENAKDIIACGFDINKTFISDLDYMGSSGFYKN 203
DB 139 VIELTDEKFLFKKLTINDVKNFAENKADIIVAGFDPKNTFIFSDLYMG--GAFYET 196

QY 204 VKIKQHVTFNQVKGIFGFTSDICIGISFPALQAPSPNSFPQIFRRTDIOCLIPCA 263
DB 197 VVRSRQITGSTAKAVFGFNSDICIKFHFAIQAAPSSSPFPHFNGAKDIPCLIPCA 256

QY 264 IDDPYFMTBDVAPRIGYPRKALHSTFEPALQAGQTMASDPNSSIFLDTAKQIKTK 323
DB 257 IDDPYFRCRDVADLTKSKPLHSHRFPALQGSTTMASDPTTALFMTDTNKKIKN 316

QY 324 KVNKAHAFSGGRDTEIEHROFGNCVDVSPMYLTFLEDDKLEQIRKDYSGAMLTKEL 383
DB 317 KINKVAFSGGQVSADLHRELGNPDVAYQYLSFFKDDVFLKECYDKYKSGELLSGEM 376

QY 384 KKALIEVLQPLAEHQARKEVTDEIVKEFMTPRKL 419
DB 377 KKLCTIELOEFVAKAFQERRAOVDEETLDFKMPVHKL 412

RESULT 7
C90190
tryptophanyl-tRNA synthetase (trps) [imported] - *Sulfolobus solfataricus*
C/Species: *Sulfolobus solfataricus*
C/Date: 24-May-2001 #sequence revision 24-May-2001 #text_change 15-Jun-2001
A/Accession: C90190
R/Singh, R.K.; Cofaloni, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-Y
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
submitted to Genbank, April 2001
A/Description: *Sulfolobus solfataricus* complete genome.
A/Reference number: A99139
A/Accession: C90190
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-386 <KUR>
A/Cross-references: GB:AE00641; NID:g13813608; PIDN:AAK40778.1; GSPDB:GN00155
C/Genetics:
A/Gene: trps
C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 40.4%; Score 907; DB 2; Length 386;
Best Local Similarity 48.6%; Pred. No. 6.2e-66;
Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

QY 31 TEAEDFVDPWTVQTSANG-IDYDKLIVFGSSKIDKELINRIERATGGRPHHFRGI 89
DB 6 TMDDEFTVPMVEV-----KGVADYDKLIVQFGTKITPEOLERPEKLTGKKPHLLRGA 59

QY 90 FFSHRDMNOVLDAVENKKPFYLYTGRGSSSAMVGHILPFIPTKMLQDVFNPLVLTOMT 149
DB 60 FFSHRDPLVLDYENKSGPFLYTGRRPSL-GMIIHLLPFIPTKMLQEKFNANLITET 118

QY 150 DDEKYLMD-LTLDOAYGDAVENAKDIIACGFDINKTFISDLDYMGSSGFYKNVVKIQ 208

Query Match 23.8%; Score 534.5; DB 2; Length 380;
 Best Local Similarity 35.4%; Pred. No. 1.2e-35;
 Matches 14; Conservative 76; Mismatches 140; Indels 41; Gaps 17;
 QY AEDPVDPTVOTSSAKGIDYKLIVRGSSKIDKELINIRATGQRPHNFRGIFSS 92
 DB ADGNDVTPAVESDD---LDYEKLIARFAGDELTDGRRFP-----DHPVNRGLFVA 53
 QY 93 HRDMNQVLDAVENKPPYLYTGGRSPSEAMHVGHLIFIFTKMLQDVNPLVYQMTDDE 152
 DB 54 GRVDVDFLTAGEOS---IYTGVPSSG-PHMLGAMVFYRARRQDFFGARVYVPLSDDE 108
 QY 153 KYUMKDLTLLOAVGDAVE-NAKDIIACGPDINKTFIF---SDLPYM-GMSSGFYKXVVKI 207
 DB 109 KYMFQKQTPAET-GDYIRALNRDLIANGFDELTIRYVDRDADVLKPLATAPAGDV--- 164
 QY 208 QKAVTFNQVKGIFGFTDSDCTGKISFPALQAPSPNSFPQIFRDRDIOCLIPCALDQ 267
 DB 165 -RHATLQNVG-----EPDNVGQAFYPAVQTAHL---LPQLVHG--EHETLVPIAVDQD 213
 QY 268 PYRRMTRDVAPRGYP--KRALHSTFPALQAGQOTMSASDRSSIFLTDTKQKITKY 325
 DB 214 PHVRVSDVAAKARYPVGKPGALMOPLSLAG-PGMSSS-AGVSIRLTDSPVAREKV 271
 QY 326 NKAFSGGRDTIEHRQFGNCVDVDFMYLTFLEDD-D-KLBOIRKDYSGAMLTGELK 384
 DB 272 RTIAYTCGRASVEHRKAGVPAEDVFOYLSAFPEEDDELARIERYRAGDLSGELK 331
 QY 385 KALIEVLQPLIAEHQARRKEVTEIVEFMTPRKLSFD 422
 DB 332 DLAAIRTEFLAHOARRALGD--VTEALDAFRLTD 367

RESULT 11

64476
 cryptophan-tRNA ligase (EC 6.1.1.2) - Methanococcus jannaschii
 N/Alternate names: cryptophanyl-tRNA synthetase
 C/Species: Methanococcus jannaschii
 C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 03-Jun-2002
 QY C/Accession: F64476
 R/Built: C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
 Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
 Ison, J.D.; Sedow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurest, M.A.
 Science 273, 1058-1073, 1996
 A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
 A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A/Reference number: A64300; MUID:96337999; PMID:8688087
 A/Accession: F64476
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-370 <BUIL>
 A/Cross-references: GB:U67582; GB:L77117; NID:g1592064; PID:AA899425.1; PID:g1592065;
 C/Genetics:
 A/Map position: FOR3375885-1376997
 A/Start codon: GTG
 C/Superfamily: mammalian cryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
 C/Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 18.2%; Score 409.5; DB 2; Length 370;
 Best Local Similarity 30.9%; Pred. No. 1.7e-25;
 Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;
 QY 40 PWTVOSSAGIDYKLIVRGSSKIDKELINIRATGQRPHNFRGIFSHRDNV 99
 DB 8 PW--ETPAV--IDYKTMEOFGVAPYVDVGLDER-----HFFFRNIIIGHRDERI 57
 QY 100 LDAYENKPPYLYTGGRSPSEAMHVGHLIFIFTKMLQ--DVFNVPVLVQMTDDEKYLW 156
 DB 58 VDAIKNNKEFAVNSGMPSRK-WHFGKMWVDLLKFOKXTDNINIFI-----ADLEAYWA 112
 QY 157 KDILTDQAYDAV-ENAKDIIACGFDINKTFISDLDYMGSSGFYKXVVKI-QKAVTEN 214

DB 113 RNMSFETKELALNEXITYNALGLDPEKINIVYLSQKQV-----KDLALISKRTNMS 167
 QY 215 QVAGIGFTDSDCTGKISFPALQAPSPNSFPQIFRDRD--DIOCLIPATIDDDYFPM 272
 DB 168 EMKAIYGFGETNIGVPAIVQVADIL---HQDLNENISPEKPVVAVPGIDODPHIRL 224
 QY 273 TRDVADR---IGYKRALHSTFPALQAGQOTMSASDRSSIFLTDTKQKITKYKNA 329
 DB 225 TRDIANRAKEFKTIPSSSYHRFWTGLLG--KMSSKRETAFLTDDETKVKKIFS-A 281
 QY 330 FSGGRDTIEHRQFG--NCDVDSFMYLTFLEDDDKLBOIRKDYSGAMLTGELKKA 386
 DB 282 KTGGRTELEHKKYGVPEECVYELFLY--HLITDKEIAETIYQCRSGELTGCKCKM 339
 QY 387 LIEVLQPLIAEHQARRKEVTEIVEK 411
 DB 340 AYERVVFLDKLKEKREKQAEIAYK 364

RESULT 12

669131
 cryptophan-tRNA ligase (EC 6.1.1.2) - Methanobacterium thermoautotrophicum (strain Delta
 N/Alternate names: cryptophanyl-tRNA synthetase
 C/Species: Methanobacterium thermoautotrophicum
 C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 03-Jun-2002
 QY C/Accession: E69131
 R/Smith, D.R.; Doncelte-Stamm, L.A.; DeJonghe, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qu, D.; Spatafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Ujwani, N.;
 Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi
 A/Reference number: A69000; MUID:98037514; PMID:9371463
 A/Accession: E69131
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-364 <MTH>
 A/Cross-references: GB:AE000812; GB:AE000666; NID:g2621298; PID:AA84757.1; PID:g2621303
 A/Experimental source: strain Delta H
 C/Genetics:
 A/Gene: MTH251
 A/Start codon: TTG
 C/Superfamily: mammalian cryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
 C/Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 17.7%; Score 397.5; DB 2; Length 364;
 Best Local Similarity 27.9%; Pred. No. 1.6e-24;
 Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;
 QY 38 VDPWTVOTSSAKGIDYKLIVRGSSKIDKELINIRATGQRPHNFRGIFSHRDNV 97
 DB 2 IDPW-----GSAK-LELYODLIENGVAPF-SEVLDEV-----PEPSWLMRGIIFGHRDYE 50
 QY 98 QUIDAYENKPPYLYTGGRSPSEAMHVGHLIFIFTKMLQDVNPLVYQMTDDEKYLW 157
 DB 51 RIISAMKGEDPAVNVGMPSGR-WAIGKMTVDOLRW-YDRMGAEIFIPIDMEAYSAK 108
 QY 158 DLTLDQAYDAVEN-AKDIIACGFDINK-----TFISDLDYMGSSGFYKXVVK 206
 DB 109 GVDPEFSRRIALIEYIAGYIALGLDEKONIHVYLOSEMLWEDLAYV----- 156
 QY 207 IOKHVTFNQVKGIFGFTDSDCTGKISFPALQAPSPNSFPQIFRDRDIOCLIPCALDQ 266
 DB 157 LAGKVNFELRAIYGTGSTSMHMYAPILQVSDILHPLQDLGGR---PVLVPVGPQD 213
 QY 267 DPFMTRDVAPRI---GYKRALHSTFPALQAGQOTMSASDRSSIFLTDTKQKIT 322
 DB 214 DPHIRLIRDIARFRDRYFPLPSSTYHRFMGGLTGG--KMSNRPKSAIFLSDTEEAR 271
 QY 323 TKVKNKAFSGGRDTIEHRQFGNCVDVDFMYLTFLE--DDDKLEOIRKDYSGAMLTG 381
 DB 272 AKI-RNAKTCGRTELEKQRELGCVPEECITETILYHMGSSBSRLEIYESCRNGTLMCG 330
 QY 382 ELKALIEVLQPLIAEHQARRKE 404

Db 331 ECKNNNAEFTIRKFEELSRYREK 353

RESULT 13

T43806

tryptophan-tRNA ligase (EC 6.1.1.2) [imported] - Encephalitozoon cuniculi (fragment)

C:Species: Encephalitozoon cuniculi

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 03-Jun-2002

R:Accession: T43806

R:Peptide: T43806

Mol. Biol. Evol. 15, 683-689, 1998

A:Title: Microsporidia, amitochochordal protists, possess a 70-kDa heat shock protein ger

A:Reference number: 226293; PMID:98277683; PMID:9615449

A:Accession: T43806

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-134 <PEY>

A:Cross-references: EMBL:AJ012470; PDB:CAA10034.1

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

C:Keywords: ligase

Query Match 17.2%; Score 386; DB 2; Length 134;

Best Local Similarity 51.5%; Pred. No. 3.2e-24;

Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

QY 35 EEPVDMVYQTSASAK----GIDYDKLIYVFGSSKIDKELINRIERATGQRPHNIFLRGIF 90

Db 3 EORITPMDVEVSTGEVPAIDYDKLINOGEKFNQALADRLKLSGRPAHYFFRKGIV 62

QY 91 FSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMVGHILPFIPTKMLQDVFNVLVYQMTD 150

Db 63 FHRDPNLLIDELANNRPYLYTGRGSSSKTMHIGITIFPLCKYQDAFKRLVLYQITD 122

QY 151 DEKYLMDKDLTD 162

Db 123 DEKFLMKSWRL 134

RESULT 14

E69461

tryptophanyl-tRNA synthetase (trps) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000

R:Accession: E69461

R:Kleink, H.P.; Claydon, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, E.F.; Glöck, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes, S. Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; PMID:98049343; PMID:9389475

A:Accession: E69461

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-420 <KLE>

A:Cross-references: GB:AE000986; GB:AE000782; NID:92689309; PDB:AA8954.1; PID:9264885

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 16.5%; Score 370.5; DB 2; Length 420;

Best Local Similarity 27.3%; Pred. No. 3e-22;

Matches 118; Conservative 60; Mismatches 166; Indels 89; Gaps 12;

QY 38 VDPMTVQTSASAKIDYDKLIVFGSSKIDKELINRIERATGQRPHNIFLRGIFFSHRDMN 97

Db 3 VPMVEVEGV----IDSKLIEFGMQPF--SEVLPEID-----NPHLMKRGALIFGRDYW 52

QY 98 QVLDAVENKKPFYLYTGRGSSSEAMVGHILPFIPTKMLQDVFNVLVYQMTDEKYLTK 157

Db 53 RILEAMQKPEPAVWSGFMPSG--LPHFGHKMTDELIVWQSAGKAFV--AIDMEASHSV 110

QY 158 DLTLDAYGDAVENAKDIIACGFDINKTIFSDLDYMGSSGFYKVVV--KIQKHVTFNQV 216

Db 111 GLSMKTEELGMLYKSIITIALGLREDATVYFGS-----KSHVYDLAFELSAFVNFSGL 164

QY 217 KQIFPTSDSDICGISFPALQAPSFNSFPQIFRDRDIOCLIPCAIDDDYFPMTRDV 276

Db 165 RAIVGFNSDTSIAKMFVTAIQADIL--HPQLSDFGPKRPVVVGVGADQDPHMLTDL 221

QY 277 APR----- 280

Db 222 AARISFSPFVEGVRVRSRKAELVLSLNDLEFDKIIYEEHMDIFGEAEIERAVRKI 281

QY 281 -----GVPKPALHSTFFPALQAGOTKMSASDPNSSIFLTDTAKQIKTKVNHKAFFSG 333

Db 282 EVELGFAFIPRSSSYNHFTTGLTGG--KMSSKRESYISLIDPPEBAKVMK--AFTGG 338

QY 334 RDTIEHROFGNCVDVDSFMYLTFLELD--DDKLEQIRKDYTGAMLTGELKALIEVLQ 392

Db 339 RATAEORRLGGEPDRCVFEVLYSFLIDSDELNQLIABECREGLCGCKMAAEVLK 398

QY 393 PLIAEHQARKEV 405

Db 399 SFLKEHQEKMEAV 411

RESULT 15

D72477

probable tryptophanyl-tRNA synthetase APE2461 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

R:Accession: D72477

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hakiwa, Y.; Jin-no, K.; Takahashi, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; PMID:99310339; PMID:10382966

A:Accession: D72477

A:Status: preliminary

A:Residues: 1-374 <KAN>

A:Cross-references: DDBJ:AP000064; NID:95105945; PDB:BA81476.1; PID:95106165

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2461

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 15.7%; Score 353; DB 2; Length 374;

Best Local Similarity 30.4%; Pred. No. 6.7e-21;

Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

QY 38 VDPMTVQTSASAKIDYDKLIVFGSSKIDKELINRIERATGQRPHNIFLRGIFFSHRDMN 97

Db 8 IDPW---GAVEIKDYDLRLTFRPSEVL--PLLRKAGKPSFLMRGIIFFGRDPD 61

QY 98 QVLDAVENKKPFYLYTGRGSSSEAMVGHILPFIPTKMLQ--DVFNVLVYQMTDEKYLTK 156

Db 62 KLEKAKGGRVAVLTGMPGSK--PHRHKLTVDOLYILQKNGFV--FVALADAEAAV 118

QY 157 KDLTLDAYGDAVEN--AKDIIACGFDINKT--FTFSLDYMGSSGFYKVVV--KIQKHVTFN 214

Db 119 RIRIGEEAVRIAEVYINMIALGLDPKOTERYFO---TVRGTDFYFLIQFSGKVTAA 174

QY 215 QVKGIFG--FTSDSDICGISFPALQAPSFNSFPQIFRDRDIOCLIPCAIDDDYFPMTR 273

Db 175 EMEALYGEITPAKMAVSLT---QAADILHVQLDVEGYR--HVVVPGADQDPHMLT 227

QY 274 RDVAPR-----GVPKPALHSTFFPALQAGOTKMSASDPNSSIFLTDTAKQIKTKVNHKA 329

Db 228 RDLARMAGVVLEERPASTYHKLQPLGDS--RKMSSSRPDSTIFLITDPEVARKLFR--A 284

QY 330 FSGGDTIEHRQFGNCV--DVSMYLTFFLEDDKLEQIRKDYTS---GAMLTGELTK 385

Db 285 LTGGRATAEORRLGGEPDRCVFEVLYSFLIDSDELNQLIABECREGLCGCKMAAEVLK 398

Fri Jul 11 11:50:10 2003

us-09-813-718-10_copy_48_471.rpr

Page 8

QY 386 ALIEVLOPLIAEHQARRKEVTDEIVKEFMTPR 417
| : ||| : | : |
DB 342 IAMEKLERFLAEHOSRLERAKTIAMKLVPEPR 373

Search completed: July 10, 2003, 12:32:17
Job time : 18.4461 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:26:29 ; Search time 9.25091 Seconds

(without alignments)
1900.998 Million cell updates/sec

Title: US-09-813-718-10_COPY_48_471

Sequence: 1 MSYKAAAGEDYKADCPGPNP.....VTDELVKEMPRKISFPDQ 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SWISSProt_40.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	99.3	471	1	SYN_HUMAN
2	2134.5	95.0	475	1	SYN_BOVIN
3	2056	91.5	475	1	SYN_RABIT
4	2032	90.5	481	1	SYN_MOUSE
5	1210	53.9	395	1	SYN_SCHPO
6	1163	51.8	432	1	SYN_YEAST
7	910.5	40.5	381	1	SYN_SULTO
8	905	40.3	380	1	SYN_SULSO
9	803	35.8	385	1	SYN_PYRAB
10	796	35.4	386	1	SYN_PYRHO
11	735	32.7	375	1	SYN_PYRAE
12	534.5	22.8	380	1	SYN_HALNI
13	409.5	18.2	370	1	SYN_MENJA
14	397.5	17.7	364	1	SYN_MERTH
15	386	17.2	134	1	SYN_ENCCU
16	370.5	16.5	420	1	SYN_ARCFU
17	353	15.7	374	1	SYN_AERPE
18	299.5	13.3	426	1	SYN_THEVO
19	268.5	12.0	426	1	SYN_TREAC
20	192	8.5	323	1	SYN_ARCFU
21	185.5	8.3	341	1	SYN_STRPN
22	182.5	8.1	340	1	SYN_STRPY
23	175.5	7.8	341	1	SYN_CLOLO
24	174.5	7.8	341	1	SYN_LACLA
25	172	7.7	394	1	SYN_YEAST
26	170.5	7.6	366	1	SYN_SULSO
27	169	7.5	395	1	SYN_AQUAR
28	166.5	7.4	346	1	SYN_CHLTR
29	163.5	7.3	353	1	SYN_BOBBU
30	159.5	7.1	337	1	SYN_TREPA
31	158	7.0	346	1	SYN_CHLPU
32	152	6.8	344	1	SYN_CHLPN
33	146	6.5	326	1	SYN_HELPY

34	138	6.1	319	1	SYN_MERTH	027795 methanobact
35	136.5	6.0	328	1	SYN_BOVIN	029465 bos taurus
36	135.5	6.0	328	1	SYN_THEMA	094952 thermocoga
37	133	5.9	328	1	SYN_BACST	P00953 bacillus st
38	132.5	5.9	343	1	SYN_MYCLE	Q49901 mycobacteri
39	132.5	5.9	528	1	SYN_HUMAN	P54577 homo sapien
40	130	5.8	326	1	SYN_HELPJ	Q921x4 helicobacte
41	129	5.7	401	1	SYN_SCHPO	O14055 schizosacch
42	128.5	5.7	379	1	SYN_YEAST	P04803 saccharomyc
43	128	5.7	334	1	SYN_HAEIN	P43835 haemophilus
44	124.5	5.5	347	1	SYN_MYCCE	P47372 mycoplasma
45	122	5.4	350	1	SYN_CAEEL	P46579 caenorhabdi

ALIGNMENTS

RESULT 1	ID	SYN_HUMAN	STANDARD	PRT	471 AA.
AC	P23381	P78535; Q9UDJ3			
DT	01-NOV-1991	(Rel. 20, Created)			
DT	01-DEC-1992	(Rel. 24, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)				
DE	(TrpRS) (TFP53) (hmRS).				
GN	WARS.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92105071; PubMed=1761529;				
RA	Rubin B.Y., Anderson S.L., Xing L., Powell R.J., Tate W.P.;				
RT	"Interferon induces tryptophanyl-tRNA synthetase expression in human				
RT	fibroblasts.";				
RL	J. Biol. Chem. 266:24245-24248(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92107982; PubMed=1763065;				
RA	Flechner J., Rasmussen H.H., Justesen J.;				
RT	"Human interferon gamma potentially induces the synthesis of a 55-kDa				
RT	protein (gamma 2) highly homologous to rabbit peptide chain release				
RT	factor and bovine tryptophanyl-tRNA synthetase.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 88:11520-11524(1991).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92112058; PubMed=1765274;				
RA	Frolova L.Y., Sudomoina M.A., Gligorieva A.Y., Zinovleva O.L.,				
RT	Kiselev L.L.;				
RT	"Cloning and nucleotide sequence of the structural gene encoding for				
RT	human tryptophanyl-tRNA synthetase.";				
RL	Gene 109:291-296(1991).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92164636; PubMed=1537332;				
RA	Buitt U., Flohr T., Boettger E.C.;				
RT	"Molecular cloning and characterization of an interferon induced				
RT	human cDNA with sequence homology to a mammalian peptide chain				
RT	release factor.";				
RL	EMBO J. 11:485-496(1992).				
RN	[5]				
RP	SEQUENCE OF 1-13 FROM N.A.				
RX	MEDLINE=96319994; PubMed=8724762;				
RA	Sokolova I.V., Narovlyanskii A.N., Amchenkova A.M., Turpaev K.T.;				
RT	"Alternative splicing of 5'-terminal exons of the human tryptophanyl-				
RT	tRNA synthetase gene.";				
RL	Mol. Biol. (Mosk) 30:319-329(1996).				
RN	[6]				
RP	SEQUENCE OF 1-141 AND 182-471 FROM N.A.				
RC	TISSUE=Spem;				
RX	MEDLINE=93292992; PubMed=7685728;				

RA Frolova L.Y., Grigorieva A.Y., Sudomoina M.A., Kiselev L.L.
 RT "The human gene encoding tryptophanyl-tRNA synthetase: interferon-
 RL response elements and exon-intron organization.";
 RN Gene 128:237-245 (1993).
 RP SEQUENCE OF 265-276; 278-296; 299-317 AND 350-365.
 RC TISSUE=keratinocytes;
 RX MEDLINE=93162043; Pubmed=1286667;
 RA Rasmussen H.H., van Damme J., Puye M., Gesser B., Celis J.E.,
 RA Vandekerckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel
 RT protein database of normal human epidermal keratinocytes.";
 RL Electrophoresis 13:960-969 (1992).
 RN [8]
 RP FUNCTION
 RX MEDLINE=92225128; Pubmed=1373391;
 RA Bange F.-C., Flohr T., Buwitt U., Boettger E.C.;
 RT "An interferon-induced protein with release factor activity is a
 RT tryptophanyl-tRNA synthetase.";
 RL FEBS Lett. 300:163-166 (1992).
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- INDUCTION: BY INTERFERON GAMMA.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M77804; AAA67324.1; -
 DR EMBL; X59892; CAA42545.1; -
 DR EMBL; M61715; AAA61298.1; -
 DR EMBL; X62570; CAA44450.1; -
 DR EMBL; S82905; AAB39381.1; -
 DR EMBL; X67920; CAB94198.1; -
 DR EMBL; X67921; CAB94198.1; JOINED.
 DR EMBL; X67922; CAB94198.1; JOINED.
 DR EMBL; X67923; CAB94199.1; -
 DR EMBL; X67924; CAB94199.1; JOINED.
 DR EMBL; X67925; CAB94199.1; JOINED.
 DR EMBL; X67926; CAB94199.1; JOINED.
 DR EMBL; X67927; CAB94199.1; JOINED.
 DR EMBL; X67928; CAB94199.1; JOINED.
 DR PIR; A41706; A41706.
 DR PIR; A41633; A41633.
 DR PIR; JH0533; JH0533.
 DR PIR; S19246; S19246.
 DR Aarhus/Ghent-2DPAGE; 3524; IEF.
 DR PHCI-2DPAGE; P23381; -
 DR Genew; HGNC; 12729; WARS.
 DR MIM; 191050; -
 DR InterPro; IPR002306; Trp tRNA-synt_1b.
 DR InterPro; IPR000738; WHEP-TRS.
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR Pfam; PF00458; WHEP-TRS; 1.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PRO1039; TRNASYNTTRP.
 DR TIGRfam; TIGR00233; trps; 1.
 DR PROSITE; PS00178; AA tRNA_LIGASE_I; 1.
 DR PROSITE; PS00762; WHEP-TRS; 1.
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT DOMAIN 19 64 WHEP-TRS.
 FT SITE 164 173 "HIGH" REGION.
 FT SITE 349 353 "KMSKS" REGION.
 FT CONFLICT 213 214 SY -> GD (IN REF. 3).
 FT CONFLICT 424 424 A -> R (IN REF. 4).

SO SEQUENCE 471 AA; 53165 MW; E96344449053A0D0 CRC64;
 Query Match 99.3%; Score 2231; DB 1; Length 471;
 Best Local Similarity 99.5%; Pred. No. 6, 9e-177;
 Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MSYKAAAGEDYKADCCPGNAPATSNHGPDAATEAEEDFPVPMWYQTSASAKIDVDKILVR 60
 DB 48 MSYKAAAGEDYKADCCPGNAPATSNHGPDAATEAEEDFPVPMWYQTSASAKIDVDKILVR 107
 QY 61 GSSKIDKEILNRIERATGQRPHHFLRGIFFSHRDMQVLDAYENKPPFLYTGRCPSSE 120
 DB 108 GSSKIDKEILNRIERATGQRPHHFLRGIFFSHRDMQVLDAYENKPPFLYTGRCPSSE 167
 QY 121 AMVGHILIPITPKMLQDVFNVLVYQMTDDKYLKMDLTDOAYGVANMDIACGF 180
 DB 168 AMVGHILIPITPKMLQDVFNVLVYQMTDDKYLKMDLTDOAYGVANMDIACGF 227
 QY 181 DIKRTIFSDLDYMGSSGPKVVKIQRHVPNOYKGIFFGFDSDICIGISPPAIQAP 240
 DB 228 DIKRTIFSDLDYMGSSGPKVVKIQRHVPNOYKGIFFGFDSDICIGISPPAIQAP 287
 QY 241 SFNSPFOQIFRDRTDIQCILPCAIDDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAQ 300
 DB 288 SFNSPFOQIFRDRTDIQCILPCAIDDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAQ 347
 QY 301 TKMSASDPNSSFLLTPTAQIKTKVKNKAFSGGRDTIEHRQGCNDVDSFMYLTFPL 360
 DB 348 TKMSASDPNSSFLLTPTAQIKTKVKNKAFSGGRDTIEHRQGCNDVDSFMYLTFPL 407
 QY 361 EDDDKLEQIRKQYTSAMTGLKALIEVLOPLIAEHQARRKEVDEIVKEEMTPRKL 420
 DB 408 EDDDKLEQIRKQYTSAMTGLKALIEVLOPLIAEHQARRKEVDEIVKEEMTPRKL 467
 QY 421 FDFQ 424
 DB 468 FDFQ 471
 RESULT 2
 ID SYM BOVIN STANDARD; PRT; 475 AA.
 AC P17248;
 DT 01-AUG-1990 (Rel. 15, Created)
 RT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)
 DE (TrpRS).
 GN WARS.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Retina;
 RX MEDLINE=91329348; Pubmed=1907847;
 RA Garret M., Pajot B., Trezeguet V., Labouesse J., Merle M.,
 RA Gandar J.-C., Benedetto J.-P., Sallatranque M.-L., Alterio J.,
 RA Gueguen M., Sarger C., Labouesse B., Bonnet J.;
 RT "A mammalian tryptophanyl-tRNA synthetase shows little homology to
 RT prokaryotic synthetases but near identity with mammalian peptide
 RT chain release factor.";
 RL Biochemistry 30:7809-7817 (1991).
 RN [2]
 RP SEQUENCE OF 17-475 FROM N.A.
 RC TISSUE=pancreas;
 RA Garret M., Trezeguet V., Pajot B., Gandar J.-C., Merle M.,
 RA Gueguen M., Benedetto J.-P., Sarger C., Alterio J., Labouesse B.,
 RA Labouesse J., Bonnet J.;
 RL Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +

diphosphate + L-tryptophanyl-tRNA (Trp).
 CC - SUBUNIT: HOMODIMER.
 CC - SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC - SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X53918; CAA37872.1; -
 DR EMBL: X52113; CAA36356.1; -
 DR PIR: A40279; YWBO.
 DR PIR: S14540; S14540.
 DR InterPro: IPR002306; Trp_cRNA-synt_1b.
 DR InterPro: IPR000738; WHEP-TRS.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR Pfam: PF00458; WHEP-TRS; 1.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 DR TIGRFAMs: TIGR00233; trps; 1.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.
 DR PROSITE: PS00762; WHEP-TRS; 1.
 DR AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 KM Aminoacyl-tRNA synthetase; WHEP-TRS.
 FT DOMAIN 24 69
 FT SITE 169 178
 FT SITE 353 357
 FT CONFLICT 17 17 L -> M (IN REF. 2).
 SQ SEQUENCE 475 AA; 53729 MW; F7E531750137E332 CRC64;

Query Match 95.0%; Score 2134.5; DB 1; Length 475;
 Best Local Similarity 95.0%; Pred. No. 6.7e-169;
 Matches 402; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

2 SYKAANGDYKADCPGPNAPPTSNHCPDTEABEDFVDPWTYQTSASAKGIDYDKLIVRG 61
 54 SYKAATGEBDYKADCPGPNAPPTSNHCPDTEABEDFVDPWTYQTSASAKGIDYDKLIVRG 113
 62 SSKIDKELINRIERATGQRPHHFLRGIFFSHRDNMQVLDAYENKKPFYLYTGRGSSSA 121
 114 SSKIDKELINRIERATGQRPHHFLRGIFFSHRDNMQVLDAYENKKPFYLYTGRGSSSA 173
 122 MHVGHILPFTKMLQDVFNVPVLIQMTDDEKYLMDLTLDQAYGDAVENANDIACGPD 181
 174 MHVGHILPFTKMLQDVFNVPVLIQMTDDEKYLMDLTLDQAYGDAVENAND-ITCGPD 232
 182 INKTFFPSLDYWGMSGGFYKVVYKIQKAVTNQVKGIFGFTSDSGICISPAIOAPS 241
 233 INKTFFPSLDYWGMSGGFYKVVYKIQKAVTNQVKGIFGFTSDSGICISPAIOAPS 292
 242 FSNSPFOIRDRDPTDCLIPCALIDDPYFRMTDVAIPRGYPPALHSTFPALAGAPT 301
 293 FSNSPFOIRDRDPTDCLIPCALIDDPYFRMTDVAIPRGYPPALHSTFPALAGAPT 352
 302 KMSASDPNSSIFLTDPAKQIKTKVNGHAFSGGRTDIEHRQFGNCDVDSFMYLTFPLE 361
 353 KMSASDPNSSIFLTDPAKQIKTKVNGHAFSGGRTDIEHRQFGNCDVDSFMYLTFPLE 412
 362 DDDKLEQIRKYDTSGAMLTGELKKALIEVLOPIAHQARRKREVDIEYKEMTPKLSF 421
 413 DDDKLEQIRKYDTSGAMLTGELKKALIEVLOPIAHQARRKREVDIEYKEMTPKLSF 472
 422 DFO 424
 473 DFO 475

RESULT 3

SW_RABIT
 ID SYW_RABIT STANDARD; PRT; 475 AA.
 AC P25512; Q28607; (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
 DE (TrpRS).
 GN MARS.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90239043; PubMed=2185472;
 RA Lee C.C., Craigen W.J., Muzny D.M., Harlow E., Caskey C.T.;
 RT "Cloning and expression of a mammalian peptide chain release factor
 RT with sequence similarity to tryptophanyl-tRNA synthetases.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3508-3512(1990).
 RN [2]
 RP REVISIONS TO 169-174 AND 227-228, AND FUNCTION.
 RX MEDLINE=94009008; PubMed=8404867;
 RA Frolova L.Y., Dalphin M.E., Justesen J., Powell R.J., Drugeon G.,
 RA McCaughan K.K., Kisselev L.L., Tate W.P., Haenni A.-L.;
 RT "Mammalian polypeptide chain release factor and tryptophanyl-tRNA
 RT synthetase are distinct proteins.";
 RL EMBO J. 12:4013-4019(1993).
 CC - CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC - SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC - SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC - SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.
 CC - CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE AN EUKARYOTIC
 CC RELEASE FACTOR (ERF).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M33460; AAA31246.1; ALT_SEQ.
 DR EMBL: U02595; AAA60257.1; -
 DR PIR: A35904; YMRBPR.
 DR InterPro: IPR002306; Trp_cRNA-synt_1b.
 DR InterPro: IPR000738; WHEP-TRS.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR Pfam: PF00458; WHEP-TRS; 1.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 DR TIGRFAMs: TIGR00233; trps; 1.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.
 DR PROSITE: PS00762; WHEP-TRS; 1.
 DR AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 KM Aminoacyl-tRNA synthetase; WHEP-TRS.
 FT DOMAIN 23 68
 FT SITE 168 177
 FT SITE 353 357
 FT CONFLICT 17 17 L -> M (IN REF. 2).
 SQ SEQUENCE 475 AA; 53799 MW; 33BC9E718FF45DC4 CRC64;

Query Match 91.5%; Score 2056; DB 1; Length 475;
 Best Local Similarity 90.5%; Pred. No. 2.1e-162;
 Matches 383; Conservative 20; Mismatches 20; Indels 0; Gaps 0;

2 SYKAANGDYKADCPGPNAPPTSNHCPDTEABEDFVDPWTYQTSASAKGIDYDKLIVRG 61
 53 SYKAANGDYKADCPGPNAPPTSNHCPDTEABEDFVDPWTYQTSASAKGIDYDKLIVRG 112
 62 SSKIDKELINRIERATGQRPHHFLRGIFFSHRDNMQVLDAYENKKPFYLYTGRGSSSA 121

```

Db 113 SSKIDKELINRERATGQRPHRFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSSA 172
QY 122 MHVGHILPFIPTKMLQDVFNVLVLTQMTDEKYLWKDLTDQAYGAVERNAKDIACGFD 181
Db 173 MHVGHILPFIPTKMLQDVFNVLVLTQMTDEKYLWKDLTDQAYGAVERNAKDIACGFD 232
QY 182 INKTFISDLDYMGMSGGFYKNNVKIKQHTFNQVKIFGFTSDCIGKISFPALQAPS 241
Db 233 VNKTFISDLDYMGMSGGFYKNNVKIKQHTFNQVKIFGFTSDCIGKISFPALQAPS 292
QY 242 FSNFSPQIFRDRDIOCLIPCAIDDDPYFRMTDVAIRIGYKPKALLHSTFPALQAGT 301
Db 293 FSNFSPQIFRDRDIOCLIPCAIDDDPYFRMTDVAIRIGYKPKALLHSTFPALQAGT 352
QY 302 KMSASDNNSSIFLTDTAQIKTKVKNKAFSGGRDTIEHROFGNCVDVVSFMYLTFPLE 361
Db 353 KMSASDNNSSIFLTDTAQIKTKVKNKAFSGGRDTIEHROFGNCVDVVSFMYLTFPLE 412
QY 362 DDDKLEQIRKDYTSGLMTGELKKALILEVLOPLAEHQAARKVETDEIVKEFMTPRKLSF 421
Db 413 DDDKLEQIRKDYTSGLMTGELKKALILEVLOPLAEHQAARKVETDEIVKEFMTPRKLSF 472
QY 422 DFDQ 424
Db 473 HYQ 475 /

RESULT 4
SYM_MOUSE STANDARD; PRT; 481 AA.
ID _SYM_MOUSE
AC P32921;
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN WARS OR WRS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=95018226; PubMed=7932716;
RX Pajot B., Sarger C., Bonnet J., Garret M.;
RT "An alternative splicing modifies the C-terminal end of tryptophanyl-
RT tRNA synthetase in murine embryonic stem cells.";
RL J. Mol. Biol. 242:599-603(1994).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC short form; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: The short isoform is widely expressed, the
CC long form is found only in embryonic stem cells.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X69656; CAA49347.1; -
CC EMBL; X69657; CAA49348.1; -
CC PIR; S31461; S31461.
CC PIR; S31462; S31462.
CC MGD; MGI:104630; WarB.
CC InterPro; IPR002306; Trp_tRNA-synt_1b.
CC InterPro; IPR000738; WHEP-TRS.

```

```

DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF00458; WHEP-TRS; 1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRFAMs; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA tRNA ligase_1; 1.
DR PROSITE; PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
FT VARSPIC 476 481 MISSING (IN ISOBFORM 2).
SQ SEQUENCE 481 AA; 54282 MW; B05H452C08074F52 CRC64;

Query Match 90.5%; Score 2032; DB 1; Length 481;
Best Local Similarity 99.6%; Pred. No. 2e-160;
Matches 380; Conservative 22; Mismatches 22; Indels 0; Gaps 0;

QY 1 MSYKAAGEDYKADCEPPNAPATSNHGPDATEAEEDFVDPWPVYQSSAKGIDYDKLIVRF 60
Db 52 MSYKAAMGEYKAGCEPPNAPATSNHGPDATEAEEDFVDPWPVYQSSAKGIDYDKLIVRF 111
QY 61 GSSKIDKELINRERATGQRPHRFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSSA 120
Db 112 GSSKIDKELINRERATGQRPHRFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSSA 171
QY 172 AMHGLHVPFIPTKMLQDVFNVLVLTQMTDEKYLWKDLTDQAYGAVERNAKDIACGFD 180
Db 181 AMHGLHVPFIPTKMLQDVFNVLVLTQMTDEKYLWKDLTDQAYGAVERNAKDIACGFD 231
QY 232 DIKTRIFSDLDYMGMSGGFYKNNVKIKQHTFNQVKIFGFTSDCIGKISFPALQAPS 240
Db 241 DIKTRIFSDLDYMGMSGGFYKNNVKIKQHTFNQVKIFGFTSDCIGKISFPALQAPS 291
QY 292 SFNSFPOIFRDRDIOCLIPCAIDDDPYFRMTDVAIRIGYKPKALLHSTFPALQAGT 300
Db 351 SFNSFPOIFRDRDIOCLIPCAIDDDPYFRMTDVAIRIGYKPKALLHSTFPALQAGT 411
QY 361 EDDKLEQIRKDYTSGLMTGELKKALILEVLOPLAEHQAARKVETDEIVKEFMTPRKLS 420
Db 412 EDDKLEQIRKDYTSGLMTGELKKALILEVLOPLAEHQAARKVETDEIVKEFMTPRKLS 471
QY 421 FDFQ 424
Db 472 FHFQ 475

RESULT 5
SYM_SCHPO STANDARD; PRT; 395 AA.
ID _SYM_SCHPO
AC 009692;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA
DE ligase) (TrpRS).
GN SPAC2F7.13C
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

```


	RA	Sgours J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A., Gentile S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holtzoy S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K., James K., Jones L., Jones M., Leather S., McDonald S., Mclean J., RA Moore P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbinowitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B., Welterns I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Filicz C., Holzer E., Moesl D., Hilbert H., Borzym K., Langer I., Beck A., Lemrach H., Reinhardt R., Polt T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Putnulle B., RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lejaune V., Motier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Roche M., Galliard C., Tallada V.A., Garzon A., Rhode G., Daga R.R., Cruzado L., Jimenez U., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P., RT Nature 415:871-880(2002).
	CC	-I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP + diphosphate + L-tryptophanyl-tRNA(Trp).
	CC	-I- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
	CC	-----
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
	CC	-----
D	DR	EMBL; Z50142; CAAG9500.1; .-
D	DR	InterPro; IPRO02306; Trp_cRNA-synt_1b.
D	DR	InterPro; IPRO02305; cRNA-synt_1b.
D	DR	InterPro; IPR001412; tRNA-synt_1.
D	DR	Pfam; PF00579; tRNA-synt_1b; 1.
D	DR	PRINTS; PRO1039; TRNASYNTTRP.
D	DR	TIGRFAMS; TIGR00233; tTPS; 1.
D	DR	POSITIVE; PS00178; AA_TRNA_LIGASE_I; 1.
K	KM	Hypothetical protein; Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT	FT	SITE 91 "HIGH" REGION.
FT	FT	SITE 275 279 "KMSKS" REGION.
SQ	SQ	SEQUENCE 395 AA; 44910 MW; B656A8B76C5DF9 CRC64;
		Query March 53.9%; Score 1210; DB 1; Length 395;
		Best local similarity 59.6%; Pred. No. 1.4e-92;
		Matches 235; Conservative 60; Mismatches 91; Indels 8; Gaps 4;
QY	QY	34 EEDEVDPEWTWQTS----SAKGIDVDKLVIRGSSKIDELINLRERATGORPHNFLRGCI 89
D	D	4 EEQLVTPEPDVGVSLVDGEKKGIDVERLLIVGGTRKTITPEGLERFKLKXKHILLRRGA 63
QY	QY	90 FFSGHRDNNQVIDAYENKKPFVLTYGRGSSSEAMVGHILPIPTFTKWLDDVFNNPLVIOMT 149
D	D	64 FFSHRDPMILIDRAYEOKKPFYLTYGRGSSSMHGMHIMPFCKMLDDVFOVPRLVLOLT 123
QY	QY	150 DDEXYKLWMD-LTLQAVGDAYENMKDIIACGPDIINKPTIFSDLDVMGSSGFYKNVVIQ 208
D	D	124 DDEKFLETKQGVSLDEDQCFARENAKDIIAVGFDPKKTIIPNNSTVYG--GAPFQNVVAIA 181
QY	QY	209 KHVFNQVKGIFFGTDSDCICKISFPALQAAPSFSNSPPQIFRDRTDIOCLPCALIDDP 268
D	D	182 KCIANQSKACFGFTDSDSIDIKIHASIQAAPSFSSSPPHFNAKDIPCILPCALIDDP 241
QY	QY	269 YFNMTDVAPRICTPKPALHLSTFPALQAGAQTKASADPNSSIFLDTTAQKITKUNKH 328

Dd	242	YFRLTRVSGSLKTKKKALLHSRFPALQGGQSMKSSAKOSSAFLFMPTDPRIKINIKNIRH	301
Cy	329	AFGSGRDITIEHRFGGNCVDVSFWYLTFELEDXLEQIRKDYSGLMTGEIKALI	388
Dd	302	AFSGSGATIEIHRREKGNPDVDVAAYQLSFFLDDDEKXIKOLYNTRYKAGTSTGMKEGECI	361
Cy	369	EVLQPLIAHQARKEVTDELVEFPM-PRLSF	421
Dd	362	KLLQQFVSDFOARKSKVDTEATLDMPDGSRLIEW	395
<hr/>			
RESULT 6			
ID	SYMC_YEAST	STANDARD;	PRT; 432 AA.
AC	Q12109;		
Dt	01-NOV-1997	(Rel. 35, Created)	
Dt	01-NOV-1997	(Rel. 35, Last sequence update)	
Dt	01-NOV-1997	(Rel. 35, Last annotation update)	
De	Tryptophanyl-tRNA synthetase, cytoplasmic (EC 6.1.1.2) (Tryptophan--		
De	tRNA ligase) (TrpRS).		
OS	WRS1 OR YOL097C OR HRA32.		
NS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
CC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.		
OX	NCBI_TaxID=4932;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=96076631; PubMed=7502582;		
RT	Vandenbol M., Durand P., Portetelle D., Hilger F.;		
RT	"Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV		
RT	including the Trp-H3 retrotransposon, the sufl(+) frameshift		
RT	suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-Ia and a		
RT	delta element.";		
RL	Yeast 11:1069-1075 (1995).		
RN	[2]		
RP	FUNCTION.		
RA	MEDLINE=97197969; PubMed=9046085,		
RT	John T.R., Ghosh M., Johnson J.D.;		
RT	"Identification and expression of the Saccharomyces cerevisiae		
RT	cytoplasmic tryptophanyl-tRNA synthetase gene.";		
RL	Yeast 13:37-41 (1997).		
CC	-1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +		
CC	diphosphate + L-tryptophanyl-tRNA(Trp).		
CC	-1- SUBUNIT: HOMODIMER.		
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.		
CC	-1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaborat-		
CC	ion between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on ways		
CC	use by non-profit institutions as long as its content is in no way com-		
CC	mercialized and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; Z48149; CAAB8164.1; -		
DR	EMBL; Z74839; CAA99110.1; -		
DR	GD; S0005457; WRS1.		
DR	InterPro; IPR002306; Trp_tRNA-synt_1b.		
DR	InterPro; IPR002305; tRNA-synt_1d.		
DR	InterPro; IPR001412; tRNA-synt_1c.		
DR	PIfam; PF00579; tRNA-synt_1b; 1.		
DR	PRINTS; PRO1039; TRNASYNTHTRP.		
DR	TIGRFAMs; TIGR00233; trps; 1.		
DR	PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.		
KM	Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.		
FT	SITE	111	120
FT	STATE	295	299
FT	SEQUENCE	432 AA; 49350 MM; CA08P169737E9736 CRC64;	
<hr/>			
Query Match	51.8%; Score 1163; DB 1; Length 432;		
Best Local Similarity	54.8%; Pred. No. 1.2e-88;		
Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;			

```

QY 30 ATEAEEDFVDPWTV-----QTSSAKGIDVDKLIYRFGSSKIDKELINRIERATGQRPNH 84
DB 19 STVKEQVTPWVEGVGVDEQGRQNDYDKLIKQFETKRVNEBTLKRFQVGRREPHF 78
QY 85 LRGRIFFSHRDMQVLDAYENKKRPFLYLTGRGSSSEAMHVGLIPIFTKMLQDVFNVL 144
DB 79 LRKGLFFSEEDPFIKIDLYEQGRPFPLYTGRGSSDSMHGMIPIFTKMLQDVFNVL 138
QY 145 VIQMTDEKYLKML-DLTLDQAYGDAVENAKDIIACGPDINKTFIFSLDLYMGSSGYYK 203
DB 139 VIELTDEKFLFKHKLITINVKNFARENAKDIIAVGDPKOTFIFSLQYMG--GAFYET 196
QY 204 VVNIQKHVTFNQVKIGIFGFTSDICIGISFPALQAAFSFNSPQIFRDRTDIQCILPCA 263
DB 197 VVAVSQITSTAAVAFGNFSDICIGIFHPSIQIATAFSSFPNVGLPDKTPCLPCA 256
QY 264 IDDDPFRMTRDVAFRIGYKPKPALHSTFPALQAAQTKMSASDPNSIFLDTAKOIKT 323
DB 257 IDDDPFRVGRVDVADKTKYKSKPALHSTRFPALQSGTTKMSADDTALFMTDIPKQIK 316
QY 324 KVKKHAFFSGRDTIEEHROFGNCDVVSFMYLTFFLEDDDKLEQIRKDYTGSMLTGEL 383
DB 317 KIMKYAFSGQVADLRELGNGPDVAVOYLSFFRDDVFLKCYDYKSGELLGEM 376
QY 384 KKALIEUPLIAEHQARRKEVDEIVKEFMTPRKL 419
DB 377 KRCIETLQEFVKAFOERRAQVDETLDKFMPVPHKL 412

RESULT 7
SYM_SULTO
ID _SYM_SULTO STANDARD; PRT; 381 AA.
AC Q976M1;
DC 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TPRS).
GN TRPS OR ST0169.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Ohnita T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Rep. 8:123-140(2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-tRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; AP000981; BAB65126.1; ALT_INIT.
CC DR Interpro; IPR002306; Trp tRNA-synt_1b.
CC DR Interpro; IPR002305; tRNA-synt_1b.

```

```

DR Pfam; PF00579; tRNA-synt_1b; 1.
DR TIGRFRAME; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA tRNA_LIGASE 1; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 82 90 "HIGH" REGION.
FT SITE 254 258 "KMSKS" REGION.
SQ SEQUENCE 381 AA; 44718 MW; 109F5A56AF7D159 CRC64;

Query Match 40.5%; Score 910.5; DB 1; Length 381;
Best Local Similarity 50.5%; Pred. No. 7.9e-68;
Matches 189; Conservative 63; Mismatches 99; Indels 23; Gaps 9;

QY 35 EDF-VDPWTVQTSASAG-IDYDKLIYRFGSSKIDKELINRIERATGQRPNHFRGIFPS 92
DB 3 QDFNVTPMEV-----KGKVDYDKLIYQFGTKITSLKRTKISINDELHWMRLRRVPS 57
QY 93 HRDMQVLDAYENKKRPFLYLTGRGSSSEAMHVGLIPIFTKMLQDVFNVLQMTDDE 152
DB 58 HRDLVDLVMDYQDGKGFPLYTGRAPSL-GMHIQHLIPIFTKMLQDKFNVNLYEITDDE 116
QY 153 KYIWK-DLTLDQAYGDAVENAKDIIACGPDINKTFIFSLDLYMGSSGYYKVVKIQKV 211
DB 117 KFMENPEYTLQTRQWADYNDIILIVGFNPDKTFIFQDTEYI--RMWPIAIKIAKRL 173
QY 212 TENQVIGIFGFTSDICIGISFPALQAAFSFNSPQIFRDRTDIQCILPCAIDDPYFR 271
DB 174 TRSEVATGTLDTSSNIGIITWYPAQIAPT-----MEKR---RCLIPAGIDDPYFR 223
QY 272 MTRDVAFRIGYKPKPALHSTFPALQAAQTKMSASDPNSIFLDTAKOIKTKVKNHAFS 331
DB 224 LQRDIAESIGYVQAQIHSGKFLPPLTGPBKSSSQPETAIYLTDPKVERKIMKAFS 283
QY 332 GGRDTIEEHROFGNCDVVSFMYLTFFLE-DDDKLEQIRKDYTGSMLTGELKKLIEV 390
DB 284 GGQPTIELRRKKGNDIVDSFQWLYMFEPPDNKIKLIEDYRSGALLTGEIKQTLIER 343
QY 391 LQPLIAHQARRKE 404
DB 344 LNDPFEHROKREE 357

RESULT 8
SYM_SULTO
ID _SYM_SULTO STANDARD; PRT; 380 AA.
AC Q97ZX0;
DC 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TPRS).
GN TRPS OR SS00452.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332256; PubMed=11427726;
RA She Q., Singh R.K., Contaloni F., Zivanovic Y., Allard G.,
RA Aweyer M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moers A., Erauno G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet W., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-tRNA SYNTHETASE FAMILY.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: A006677; AAK40778.1; ALT INT.
 DR Interpro: IPR002306; Trp_cRNA-synt_1b.
 DR Interpro: IPR002305; tRNA-synt_1b.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PRO1039; TRNASYNTHTRP.
 DR TIGRPFAM: TIGR00233; trps; 1.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 FT SITE 81 89 "HIGH" REGION.
 FT SITE 253 257 "KMSKS" REGION.
 SQ SEQUENCE 380 AA; 44691 MW; CF8344CF63883680 CRC64;

Query Match 40.3%; Score 905; DB 1; Length 380;
 Best Local Similarity 48.7%; Pred. No. 2.2e-67;
 Matches 190; Conservative 68; Mismatches 108; Indels 24; Gaps 10;

QY 34 EEDFVDPMVYQTSAGK-IDYDKLIYRFGSSKIDKELINRERATGQRPHHFLRGIFPS 92
 DB 3 DEFTVPMVEV-----KQKVDYDKLIYQFTQKITEELKORIKVLAIDL-HVMLRRNVFPS 56

QY 93 HRDMNOVLDAVENKKEFFLYLTGRGSSSEAMHVGHLIPFTFKLQDVENVPLVIQMTDDE 152
 DB 57 HRDLDLVANDYKSKFFLYLTGRAPSL-GMHLGHLIPFTFKLQKFRANLYITETDDE 115

QY 153 KILWK-DLTLDQAYGDAVENAKDIIACGPDINKTFIFSLDLYMGSSGFYKVVKIQKEV 211
 DB 116 KVMRNEEFITLDQTRSAVAYNIIIDIIAVGFNPDKTFIFQDEYI---RNMVPIIVKIAKCL 172

QY 212 TENVVVGIGFTDSDCIKISFPALQAPSFNSFPQIFRDRDIOCLIPCAIDQDPYFR 271
 DB 173 TSEVAVATGGLDASSNIGLIFPALDIAPL-----MEKK---RCLIPAGIDQDPYR 222

QY 272 MTRDVAPRIGYPPKALLHSTFPALQAOAKSPNSIFLTDPAKQIKTKVNGHAPS 331
 DB 223 LQDIDIESLGYYAAQIHSTFLPPLTGPEGKSSNPERAIYIVDPKVERKIMKTAAS 282

QY 332 GGRDTIEHRQFGNCDDVVSFMYLTFLEDDD-KLEQIRKDYTSGLMGLGELKALIEV 390
 DB 283 GGQPTIELHRKYGNDPEIDVPFQMLYFPFEDDNRIKEIEEYRSGLMGLGELKALIDK 342

QY 391 LQPLIAEHQARKKEVTDVYKEMTPRKIS 420
 DB 343 LNNFLEHR-RRREAKELVHVKKYDGKLA 371

RESULT 9
 ID SYM PYRAB STANDARD; PRT; 385 AA.
 AC Q9UY11;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
 DE (TrpRS).
 GN TRPS OR PAB1111.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_Taxid=29292;
 RN SQUENCE FROM N.A.
 RP STRAIN=G5 / Orsay;
 RC Heilig R.;

RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
 RT structure and evolution";
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC dihydrophate + L-tryptophanyl-tRNA(Trp).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AJ248288; CAB50601.1; -
 DR Interpro: IPR002306; Trp_cRNA-synt_1b.
 DR Interpro: IPR002305; tRNA-synt_1b.
 DR Interpro: IPR001412; tRNA-synt_1.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PRO1039; TRNASYNTHTRP.
 DR TIGRPFAM: TIGR00233; trps; 1.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 FT SITE 82 90 "HIGH" REGION.
 FT SITE 253 257 "KMSKS" REGION.
 SQ SEQUENCE 385 AA; 45100 MW; 4C29D0141497B12 CRC64;

Query Match 35.8%; Score 803; DB 1; Length 385;
 Best Local Similarity 45.3%; Pred. No. 6.2e-59;
 Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

QY 35 EDF-VDPMTVQTSAGKIDYDKLIYRFGSSKIDKELINRERATGQRPHHFLRGIFPSH 93
 DB 3 EDFKVTPEVEGV-----VDYNNLIHFGSPITEELKETAELTSELPLFFRRKFFPSH 58

QY 94 RDMNOVLDAVENKKEFFLYLTGRGSSSEAMHVGHLIPFTFKLQDVENVPLVIQMTDDEK 153
 DB 59 RYDVKVLDYBERGRFFLYLTGRGSG-PMHIGHIIPFATKMLQKFGVNLVIQITDDEK 117

QY 154 YLMKD-LTLDQAYGDAVENAKDIIACGPDINKTFIFSLDLYMGSSGFYKVVKIQKHYT 212
 DB 118 FLFKENLTFEDTKHVAVENIIDLIVGFPDPTFLFQNSEF---TKIYENAIPIAKKIN 173

QY 213 FNOVVGIGFTDSDCIKISFPALQAPSFNSFPQIFRDRDIOCLIPCAIDQDPYFRM 272
 DB 174 FEMAAVAVGFTIEQSNIGMIFFPALDIAPL-----FEKR---RCLIPALDQDPYRRL 223

QY 273 TRDVAPRIGYPPKALLHSTFPALQAOAKSPNSIFLTDPAKQIKTKVNGHAPS 332
 DB 224 QRDFAESIGYYTAAIHSTFVPSLTSGLKMSASRPERAIYITDSDPEVEKKVMKFAALTG 283

QY 333 GGRDTIEHRQFGNCDDVVSFMYLTFLEDDDKLEQIRKDY---TSGMGLGELKALIE 389
 DB 284 GRPITLKEQREKGEPEKCVFEMLEIFFEEDK-KLERYIACKXNGELTGECRRYILS 341

QY 390 VLQPLIAEHQARKKEVTDVYK 411
 DB 342 KIQEFLEKQKRRKKAQKQIEK 363

RESULT 10
 ID SYM PYRHO STANDARD; PRT; 386 AA.
 AC Q9J584;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
 DE (TrpRS).
 GN TRPS OR PH1921.

OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kousugi H., Hosoyama A., Nagai Y.,
 RA Sekai M., Ogura K., Otsuka R., Nakazawa H., Takamaya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shitaya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Ref. 5:55-76(1998).
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; AP000007; BAA31046.1; ALT-INIT.
 DR InterPro; IPR002306; Trp-tryptophan-synt_1b.
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PR01039; TRNASYNTHTRP.
 DR PROSITE; PS00178; AA-TRNA-LIGASE_I; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KM Complete proteome.
 FT SITE 82 90 "HIGH" REGION.
 FT SITE 253 257 "KMSKS" REGION.
 SQ SEQUENCE 386 AA; 45305 MW; 9E3C392F4028B2D CRC64;
 Query Match 35.4%; Score 796; DB 1; Length 386;
 Best Local Similarity 44.0%; Pred. No. 2,4e-58;
 Matches 172; Conservative 68; Mismatches 125; Indels 26; Gaps 9;
 QY 34 EEDFVDPMTVOTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHFLRGIFPSH 93
 DB 3 EEFRTVPEVAGV-----VDYDKLKHGTSPLTEDLEKTAELTKSELPIFFRRKFFPSH 58
 QY 94 RDMNQVLDAYENKKPFYLYTGRGSSSEAMVGHILPFTFKWLDVFNVLVIQMTDDEK 153
 DB 59 RVDLDLIXQVEBGRGFLYGRGPGS--PMHIGHILFFPAKMOEKRGVLYIQITDDEK 117
 QY 154 YLMKD-LTLDQAVGDAVENAKDIIACGFDINKTFIPSDLDYMGSSGFFYKNVVKIQKAVT 212
 DB 118 FLTKENLTFTDTRKWAYNDILDIIVAGFDPDKTFIFONSEF-----TKIYEMALPIAKIN 173
 QY 213 FNVKGIFFGFTSDDCIGKISFPALIOAAPSNSFPQIFRRTDIQCLIPALIDDPYFRM 272
 DB 174 FSNAAKVFGEFTESKIGMIFFPALIOAPTF-----FEKK--RCILIPALIDDPYWRLL 223
 QY 273 TRDVAPRIGYKPKPALHSTFFPALQGAOTKMSADPNSSIFLTDJAKQITKTKNKAHPSG 332
 DB 224 QRFASLSIGYKTKALHSHKVPSTLSGKMSKSPETALTYLSDPEDVAKYKWKFTLVG 283
 QY 333 GRDTIEBHPGFGNCDDVVSFMVLTFFLEDDEKLEQIRKDY---TSGAMLTGELKALIE 389
 DB 284 GRPTLKEQRKGEPEKCVVFKWLETFEEDDK--KLKERYVACRKGELTGCECKRYLIS 341
 QY 390 VLOPLIAEHQARKKVTDEIVKKEFMPRKLS 420
 DB 342 KIOEFLKEHORRRK--AEKLVKFKYTKGLA 371

RESULT 11
 SYM PYRAE
 ID SYM PYRAE STANDARD, PRT, 375 AA.
 AC 082T05;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
 DE (trpS).
 GN TRPS OR PAE3091.
 OS Pyrococcus aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Thermoproteaceae; Pyrocaculum.
 OX NCBI_TaxID=13773;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX PubMed=11792869;
 RA Fitz-Gibbon S.T., Lader H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrocaculum
 RT aerophilum.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC
 CC -----
 CC EMBL; AE009911; AAL64664.1; -.
 DR InterPro; IPR002306; Trp-tryptophan-synt_1b.
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR TIGRPFAM; TIGR00233; trpS; 1.
 DR PROSITE; PS00178; AA-TRNA-LIGASE_I; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KM Complete proteome.
 FT SITE 81 89 "HIGH" REGION.
 FT SITE 258 262 "KMSKS" REGION.
 SQ SEQUENCE 375 AA; 43178 MW; 3DDFB85DA680F116 CRC64;
 Query Match 32.7%; Score 735; DB 1; Length 375;
 Best Local Similarity 41.6%; Pred. No. 2.5e-53;
 Matches 162; Conservative 79; Mismatches 130; Indels 18; Gaps 11;
 QY 34 EEDF-VDPMTVOTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHFLRGIFPS 92
 DB 2 EEFVVTPEVAGR-----VDYDKLKHGKAPKLDKDEVALLKAYGE--VHPLIRRGFVYA 56
 QY 93 HDMNQVLDAYENKKPFYLYTGRGSSSEAMVGHILPFTFKWLDVFNVLVIQMTDDE 152
 DB 57 HHDFFPIMKHQBEGRWALYTTGRGPGS--PVHIGHWVWLLKMFSDKFLLEVYFQITDDE 115
 QY 153 K-YLMKD-LTLDQAVGDAVENAKDIIACGFDINKTFIPSDLDYMGSSGFFYKNVVKIQKAV 211
 DB 116 KYDDPEPMLEATWATWYNALDVALGFSPEKHLIIDTKI---KPIYPIAVNAVAKU 172
 QY 212 TENQVKGIFGFTSDDCIGKISFPALIOAAPSNSFPQIFRRTDIQCLIPALIDDPYFR 271
 DB 173 TNNVYATGFTDSITNIGLIFYPSLIQIAVAFPT--ELRREATPV--LIPALIDDPYFR 228
 QY 272 MTRDVAPRIGYKPKPALHSTFFPALQGAOTKMSADPNSSIFLTDJAKQITKTKNKAHPS 331
 DB 229 LARDIADALGYKPKPSTLYSKFTMALTG--ESKMSASNPDSATYITLDEKTVRRKV--MNAFT 286

```

QY 332 GGRDTEIEHROFGNCDVVSFWYLTFFLEDDDKLEQIRKDYTSGAMLTGELKALILEV 391
DB 287 GGRPTAEBOCKTYGANEVPCVPHYHMLFPDDASVEKTRKODCKSGALCGECKLKHKEKI 346
QY 392 OPLIAHQARKEVTDIEVKEFMTPRKLS 420
DB 347 TFLKEHRRERREKARKG-VDEYRLSYKLS 374

RESULT 12
SYM_HALN1
ID SYM_HALN1 STANDARD; PRT; 380 AA.
AC 09HN66;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE TrpCophany1-cRNA synthetase (EC 6.1.1.2) (TrpCophan--cRNA ligase)
DE (TrpRS).
OS TRPS OR TRPS2 OR VNGS232G.
OC Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shultz H.D., Lasky S.R., Balliga N.S., Thorson V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,
RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddock D.G., Jablonki P.E., Krebs M.P., Angerling C.M., Dale H.,
RA Isebnager T.A., Beck R.F., Pohlshroder M., Spudich J.L., Jung K.-H.,
RA Alam W., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Jiang P., Riley M., Hood L., DasSarma S.;
RA "Genome sequence of Halobacterium species NRC-1."
RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RT -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + cRNA(Trp) = AMP +
RT -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AE05109; AAG20355.1; -
CC InterPro: IPR002306; Trp_cRNA-synt_1b.
CC InterPro: IPR002305; cRNA-synt_1b.
CC InterPro: IPR001412; cRNA-synt_1.
CC Pfam: PF00579; cRNA-synt_1b; 1.
CC PRINTS: PR01039; TRNASYNTHRP.
CC TIGR: TIGR00233; crds; 1.
CC PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
CC Aminoacyl-cRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC FT SITE 74 82 "HIGH" REGION.
CC SITE 249 253 "KMSKS" REGION.
CC SEQUENCE 380 AA; 41936 MW; 0F7B2B95386404F CRC64;

Query Match 23.8%; Score 534.5; DB 1; Length 380;
Best Local Similarity 35.4%; Pred. No. 9.6e-37;
Matches 141; Conservative 76; Mismatches 140; Indels 41; Gaps 17;
QY 33 AEDFPDPTVTGTSKAGIDYDKLYRFGSSKIDKELNIRERATGQRPHHFLRGIFPS 92
DB 3 AGNDNTTPAVESDD---LDYKRLARFGADLTDQRRARFP-----DHPVNRGLFYA 53
QY 93 HNDMNQVLDAVENKPKFVLYTGSGPSSSEAMHVGHLPIFFTKMLQDVFNVLVIQWTDDE 152

```

```

DB 54 GSDVDVDFILTAGS-----IVTGVGSSG--PMELGHAMVYFARRLODEEGARVYVLSDE 108
QY 153 KTLMKOLTLDDQYGVAVE-NAKDIIACGPDINKTIF--SDLDYM-GMSGFFYNNVYKI 207
DB 109 KWFMDQTPAET-GDYLANLRDLAVGFDEDELTRIVVDTBDADVLVPLATFADV--- 164
QY 208 QKHVTFNQGKIFGSTDSCIGKISFPALQAPSPNSFPQIFNRRTIOCLIPALIOD 267
DB 165 -RHATLQNVYG---EBDNGQAFYPAVQTRHLL--LPQVHG--BHEITVPIAVDOD 213
QY 268 PYFRMTDVAPRIGYP--KPALLHSTFPALQAGTKKASAPDPSISFILTAKQIKTV 325
DB 214 PHVRYSRQVAAKARYPVGRKPGALLMQFLPISLAG-PGKXSS-AGVISRLTDSPTVREKV 271
QY 326 NKHAFSGRDTIEHROFGNCDVVSFWYLTFFLEDDDKLEQIRKDYTSGAMLTGELK 384
DB 272 RHAAYTGRASVEEHRAGVPAEDVPFOYLSAFEPDPAELARIEREYRAGDILSGELK 331
QY 385 KALIEVLOPLIAHQARKEVTDIEVKEFMTPRKLSFD 422
DB 332 DIAADRITFPLAAHQRRALGD--VTEALDAFRLTDD 367

RESULT 13
SYM_METJA
ID SYM_METJA STANDARD; PRT; 370 AA.
AC 058810;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TrpCophany1-cRNA synthetase (EC 6.1.1.2) (TrpCophan--cRNA ligase)
DE (TrpRS).
OS TRPS OR MJ1415.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayson R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Metzick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weisman J.F., Fuhrmann J.L., Nguyen D.,
RA Uiterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Frazer C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RT Science 273:1058-1073(1996).
RT -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + cRNA(Trp) = AMP +
RT -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, U67582; AAB99425.1; -
CC TIGR: MJ1415; -
CC InterPro: IPR002306; Trp_cRNA-synt_1b.
CC InterPro: IPR002305; cRNA-synt_1b.
CC InterPro: IPR001412; cRNA-synt_1.
CC Pfam: PF00579; cRNA-synt_1b; 1.
CC PRINTS: PR01039; TRNASYNTHRP.

```

TIGRNAME: TIGR00233; ttps; 1.
 DR PROSITE: PS00178; AA TRNA_LIGASE I; FALSE NEG.
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 FT SITE 75
 FT SITE 255
 SQ SEQUENCE 370 AA; 42660 MW; E6C71107CE82859D CRC64;
 Query Match 18.2%; Score 409.5; DB 1; Length 370;
 Best Local Similarity 30.9%; Pred. No. 26-26; Indels 41; Gaps 16
 Matches 119; Conservative 72; Mismatches 153;
 QY 40 PWTVTSSAKGIDYDKLIVFGSSKIDKELINIRERATGQRPHEFLRGIFSHRDNQV 99
 Db 8 PW-ETPAV-IDYKKTMEQFVKRIVDLGDLKE-----HFFRRNIIGHRDFERI 57
 QY 100 LDAIYENKKEPFIYITGRGSSSEAMRHGHLIFFTYKIQ--DVFNPLVITQNTDDEKILM 156
 Db 58 VDAIKNNNEFAVSSGMPSPGK-MHGCHKVAVDLNLFYQKYTNINIP1--ADLEAWYA 112
 QY 157 KDLTDAQYGDV-ENAKDIIACGPDINTKFIISDIDYMGSSGFYKNVKI-QKHVTEN 214
 Db 113 RNMSEETTKELALNRYITNYIALGIDPEKINLYLQSKYV-----KDLALLSKRTMS 167
 QY 215 QVKGIFGFTSDSDCIKISFPALQAPBSNSNFPQIFRDT--DIQILICAIIDQPYFM 272
 Db 168 EMKAIYGEKGETNIGHVPAPVIOVADIT--HPOLDENISPEPKPVAVPVGIDQDPHRL 224
 QY 273 TRDVNPR---IGYPPALHSTFPFPAILOGATKMSASDPNSIFLITDAKQIKTKVKNKA 329
 Db 225 TRDIANRAKEKFFETFPSSYTHHFMGGLGG--KMSSSKRETAIFLTDDEKIVYKKKIFS-A 281
 QY 330 FSGGRTIEEHRQFGS--NCDVDSFMYLTFLEBDDKLEIQRDYSGLTSGAMLTGELKA 386
 Db 282 KTGGRETEIEHHKYGVEPECVYELFTY--HLIIDDKELAEIYQCRSGELTGCKCKKM 339
 QY 387 LIEVLQPLIAEHQARKKVTDEIVK 411
 Db 340 AYERVVEFLKDLKERQAKETAIVK 364
 RESULT 14
 SYM_METH STANDARD; PRT; 364 AA.
 ID SYM_METH
 AC 026352;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TrpRS).
 DE TRPS OR MTH251.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanobacter.
 RX NCBI_TaxID=187420;
 RM [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Altdredge T., Bashlitzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Potcher B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Saefer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-T., Rice P., Noelling U., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT delta: functional analysis and comparative genomics";
 RL J. Bacteriol. 179:7135-7155 (1997).
 CC -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Trp) = AMP +
 CC dihydrophate + L-tryptophanyl-tRNA (Trp).
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -I- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

```

CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL; AB000812; BAB84757.1; -.
DR InterPro; IPR002305; TRP_cRNA-synt_1b.
DR InterPro; IPR002305; rRNA-synt_1b.
DR InterPro; IPR001412; rRNA-synt_1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PRO1039; TRNASYNTHTRP.
DR TIGRFAMs; TIGR00233; trpS; 1.
DR PROSITE; PS00178; AA_tRNA_LIGASE_I; FALSE NEG.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KM Complete proteome.
KW SITE 70 "HIGH" REGION.
FT SITE 251 255 "KMSKS" REGION.
SQ SEQUENCE 364 AA; 41301 MW; C2F34890338F61D CRC64;

Query March 17.7%; Score 397.5; DB 1; Length 364;
Best local similarity 27.9%; Pred. No. 1.9e-25;
Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;

QY 38 VDPMTVOTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHFLRGIFFSHRDM 97
Db 2 IDPW----GSAK-LEYODLIENFGVRPF-SEVLDEV-----PEPSWMRRGIIIFGHDIYE 50

QY 98 QVLDAYENKKPFYLYTGRGSSSEMHGHILPFTKMLDGVFNVPVLVIMTDEKYLK 157
Db 51 RIISAMKKGDFEFAVVTAMPPSGR-WHLGHKIVYDQLRW-YDRMGAEIIFIPADAEVYSAR 108

QY 158 DLTLDQAVGVDAVEN-AADIIAACGFDINK-----TFIFSDLDYMGSSGFFYKNVVK 206
Db 109 GVDEDSRRRIALIEYVINGYIALGLDELKDNHIVYLGSENLVVEDLAVY----- 156

QY 207 IQKHVNTNQVKGIFGFPTSDSCIGKISPAIQAPBSNSFPQIRDRDTDIOCLIPCAIDQ 266
Db 157 LAGKVNENELRAIYIGFTGSSMAHVAPIIOVSDIILHPOLDELGPR--PVIVGAPGQ 213

QY 267 DPFYRMTRDVAVPI----GYPKPALIISTFFPALQAGOTKMSASDPNSSIFLTDTAAQIK 322
Db 214 DPHRLRTRDIARRDYRGFTLPBSTYHFRMGILTG--KMSNRKPSAIFLSDTPPEAE 271

QY 323 TKVKAHPSGGSDITIEHRQFGGCDVAVSFMYITFLE-DDDKLEQIRDYSGAMLTG 381
Db 272 AKI-RNAKTGGRLETKQRELGAVPECECIYETLLYHMSGDSRLIEIYESCRNGTLTMC 330

QY 382 ELKRALIEVLOPLAEHQARKE 404
Db 331 ECKNNTAELFKPEELSVKREK 353.

RESULT 15
SYW_ENCCU STANDARD; PRT; 134 AA.
AC 096771;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS) (Fragment)
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_Taxid=6035;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277683; PubMed=9615449;
RA Peyretallade E., Brousolle V., Peyret P., Metenier G., Gouy M.,
RA Vivares C.P.
```

```

RT "Microsporidia, amitochoondrial protists, possess a 70-kDa heat shock
RT protein gene of mitochondrial evolutionary origin.";
RL Mol. Biol. Evol. 15:683-689(1998).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ012470; CAA10034.1; -.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 89 98
FT NON_TER 134 134
SQ SEQUENCE 134 AA; 15744 MW; 60E2935B7E1E344F CRC64;

Query Match 17.2%; Score 386; DB 1; Length 134;
Best Local Similarity 51.5%; Pred. No. 4,4e-25;
Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

OY 35 EDFVDPWTVTQTSASAK-----GIDYDKLIVRPSSSKIDKELINRIERATGQRPPIFLRGIF 90
DB 3 EGRITPMDEYEVUSTDEVPVPAIDYDKITINRGCEKFNQALADRIEKLSGKPAHFFRRGIY 62
OY 91 FSHRDMNQVLDAVENKKPFVLYTGRGSSSEAMAVGH.IPIFTKMLQDVFNVLVIQMTD 150
DB 63 FAHRDFNLLDDEIANNRPFLYLYGRGPPSSKTMHIGHTIPFLCKYQDAFKIRLVITQITD 122
OY 151 DEKYLKQDLTLD 162
DB 123 DEKFLKSMWLE 134

```

Search completed: July 10, 2003, 12:30:34
Job time : 10.2509 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:27:20 ; Search time 28.7806 Seconds

(without alignments)
3035.516 Million cell updates/sec

Title: US-09-813-718-10_COPY_48_471

Perfect score: 2246
Sequence: 1 MSYKAAAGDYKADCEPQNP.....VTDELYKEMPRKLSFDRQ 424

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriopl:*
- 17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2051	91.3	475	11 Q9DC65	Q9dc65 mus musculus
2	2047	91.1	481	11 Q9J558	Q9j558 mus musculus
3	1629	72.5	329	11 Q9D8R9	Q9d8r9 mus musculus
4	1537	68.4	305	11 Q70184	Q70184 cavia porce
5	1382	61.5	430	5 Q9U4Y1	Q9u4y1 drosophila
6	1379.5	61.4	420	5 Q9U4Y0	Q9u4y0 drosophila
7	1378	61.4	430	5 Q9VHG2	Q9vHG2 drosophila
8	1301	57.9	402	10 Q9SR15	Q9sR15 arabidopsis
9	1027	45.7	324	5 Q9U1R2	Q9u1r2 caenorhabdi
10	942	41.9	385	5 Q8SQY5	Q8sqY5 encephalito
11	826.5	36.8	385	17 Q8U453	Q8u453 pyrococcus
12	800.5	35.6	490	5 Q9U1F5	Q9u1f5 leishmania
13	682	30.4	136	6 Q9S295	Q9s295 sus scrofa
14	388.5	17.3	136	6 Q9T588	Q9t588 bos taurus
15	329.5	14.7	437	17 Q8TUA1	Q8tUA1 mechanosarc

17	310	13.8	111	5 Q95YL8	Q95YL8 encephalito
18	294	13.1	107	5 Q9U533	Q9u533 trypanosoma
19	286	12.7	159	4 Q9UD15	Q9ud15 homo sapien
20	274.5	12.2	513	17 Q9H83	Q9h83 halobacteri
21	191	8.5	324	17 Q8TX22	Q8tX22 methanopyru
22	188	8.4	364	17 Q9YA64	Q9yA64 aeropyrum p
23	181.5	8.1	331	17 Q97921	Q97921 thermoplas
24	172	7.7	385	10 Q8S9J2	Q8s9J2 arabidopsis
25	171.5	7.6	351	10 P93018	P93018 arabidopsis
26	167	7.4	351	16 Q9RV06	Q9rv06 deinetococcus
27	164.5	7.3	895	10 Q9SGN2	Q9sgN2 arabidopsis
28	158.5	7.1	375	17 Q9V027	Q9v027 pyrococcus
29	156	6.9	372	17 Q8ZY77	Q8zy77 pyrobaculum
30	154.5	6.8	102	1 Q07119	Q07119 halobacteri
31	153.5	6.8	317	17 Q8ST11	Q8sT11 methanosarc
32	152.5	6.8	682	5 Q9N988	Q9n988 leishmania
33	149.5	6.7	327	17 Q9HN62	Q9hN62 halobacteri
34	149.5	6.7	408	10 P93363	P93363 nicotiana t
35	148	6.6	316	17 Q8ZM77	Q8zM77 pyrobaculum
36	146	6.5	375	17 Q8S739	Q8s739 pyrococcus
37	145	6.5	332	17 Q9HKT3	Q9hKt3 thermoplas
38	145	6.5	337	16 Q8Z0T0	Q8z0T0 salmonella
39	142.5	6.3	375	17 Q8U2H3	Q8u2H3 pyrococcus
40	142.5	6.3	528	11 Q91WQ3	Q91wQ3 mus musculu
41	140	6.2	337	16 Q8ZK00	Q8zK00 salmonella
42	139	6.2	294	17 Q96VY3	Q96vY3 sulfobolus
43	135	6.0	365	16 Q9KNV7	Q9knV7 vibrio chol
44	134.5	6.0	525	5 Q9VAV60	Q9vAV60 drosophila
45	130.5	5.8	419	16 Q92BB1	Q92bB1 listeria in

ALIGNMENTS

RESULT 1	ID	Q9DC65	PRELIMINARY:	PRT:	475 AA.
AC	Q9DC65	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	Adult male lung CDNA, RIKEN full-length enriched library, clone:1200002C07, full insert sequence.				
GN	WARS				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=LUNG;				
RX	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada S.,				
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batdalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,				
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.D., Bult C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Guetincich S., Hill D., Holtman M., Hume D.A., Kamita M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Sorch K.-F.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,				
RA	Hayashizaki Y.;				
RT	"Functional annotation of a full-length mouse CDNA collection.";				
RL	Nature 409:685-690(2001).				
DR	EMBL, AK004541, BAB23357.1; -				

DR MGD; MGI:104630; Wars.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; Ttp_tRNA-synt_1b.
DR InterPro; IPR000738; WHEP-TRS.
DR Pfam; PF00579; tRNA-synt_1d; 1.
DR Pfam; PF00458; WHEP-TRS; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRFAMs; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA tRNA_LIGASE_I; 1.
DR PROSITE; PS00762; WHEP-TRS; 1.
SQ SEQUENCE 475 AA; 53641 MW; C3A67EE85521DE4C CRC64;

Query Match	Similarity	91.3%	Score 2051	DB 11	length 475
Best Local	Similarity	90.3%	Pred. No. 4,9e-166		
Matches	383	Conservative	21	Mismatches	20
				Indels	0
				Gaps	0
QY	1	MSYKAAAGEDYKADCPGNGPAPATSNHGPDATAEEDFVDDPMTVOTSSAKGIDYDKLYRF	60		
Db	52	MSYKAAAMEEYKACGPNGNPATGARNCDSDATASDPPVDPMTVOTSSAKGIDYDKLYRF	111		
QY	61	GSSKIDKELINRIEATGQRPHPFLRRGIFFSHRMNNVLDAENKKFPLYLTGGRPSSE	120		
Db	112	GSSKIDKELINRIEATGQRPHPFLRRGIFFSHRMNNVLDAENKKFPLYLTGGRPSSE	171		
QY	121	AMAVGHLLPIEFTKMLQDVFNVLVLTQMTDDEKYLWKQDITLDOAVGDAVENAKDIIACGF	180		
Db	172	AMHLGHLVPIFTKMLQDVFNVLVLTQMSDDEKYLWKQDITLDOAVSYVENAKDIIACGF	231		
QY	181	DINKTFISDIDLVMSSSGFYKNVYKIQGHVFNQVKGIFGFTSDDCIGKISFPAIOAAP	240		
Db	232	DINKTFISDIDLEVMQSGFGFIRNVYKIQGHVFNQVKGIFGFTSDDCIGKISFPAIOAAP	291		
QY	241	SFSNSFPIFDRDITDQCLIPCAIDODPYFRMTRVAPRIGYKPKALLHSTFFPALQAG	300		
Db	292	SFSNSFPIFDRDITDQCLIPCAIDODPYFRMTRVAPRIGYKPKALLHSTFFPALQAG	351		
QY	301	TKMASASDNNSSIFLDTPAKQIKTKYKNKAFSGSGRDTIEBHRPGFNCQDVSEFWLTFPL	360		
Db	352	TKMASASDNNSSIFLDTPAKQIKSKKNKAFSGSGRDTIEBHRPGFNCQDVSEFWLTFPL	411		
QY	361	EDDDKLEQIRKDYTSGAMLGELKALILEVLOPLIAEHQARRKEVTDIEVKEFWMPRKLS	420		
Db	412	EDDDKLEQIRKDYTSGAMLGELKALILEVLOPLIAEHQARRKAVTBEVKEFWMPRKLS	471		
QY	421	EDFQ 424			
Db	472	FHFQ 475			

RESULT 2			
099J58			
ID 099J58	PRELIMINARY;	PRT;	481 AA.
AC 099J58:			
DT 01-JUN-2001 (TREMblrel. 17, Created)			
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)			
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)			
DE tryptophanyl-tRNA synthetase.			
GN WARS.			
OS Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX NCBI_TaxID=10090;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC TISSUE=MAMMARY TUMOR. WAP-TAG MODEL. 5 MONTHS OLD;			
RL Straubeberg R.;			
RA Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.			
DR EMBL: BC003450; AA03450.1; -.			
DR MGD: MGI:104630; WARS.			
DR InterPro: IPR002305; tRNA-synt_1b.			
DR InterPro: IPR001412; tRNA-synt_1.			
DR InterPro: IPR002306; ttp-tRNA-synt_1b.			

Query Match	Score	DB	Length
Best Local Similarity	90.1%	Pred. No. 1,1e-165	
Matches 382; Conservative	22;	Mismatches 20;	Indels 0; Gaps 0;

QY	1	MSYKAAAGEDYKADCCPGNPNAPYTSNHPDAAEAEDEFDVDPWTYOTSXAKGIDYDKLIYRF	60
Db	52	VSYKAAAGEEYKACCPGPNFTAGNCSDSATKASEDFVDPWTVRTSSAKGIDYDKLIYQF	111
QY	61	GSSKIDKELNIRIRATGQRPNNHNLARGIFPSSHNDNQVLDAYENKKPFIYLTGRGSE	120
Db	112	GSSKIDKELNIRIRATGQRPNNHNLARGIFPSSHNDNQVLDAYENKKPFIYLTGRGSE	171
QY	121	AMHYGHIPIPIFTKMLDQVFNVPVLIQWTDDEKILMKDLTLDOAYGDAVENAKDIIACGF	180
Db	172	AMHGHVLPPIFTKMLDQVFNVPVLIQMSDDEKILMKDLTLDOAYSTYVENAKDIIACGF	231
QY	181	DINKTFIFSDLDYGMSSGFYKNVVKIQKVTENVQKIGFETDSDCIKISPALQAP	240
Db	232	DINKTFIFSDLEUYGSGPGFYRNVVKIQKHTFNQVKGIFGFTDSDCIKISPAYQAP	291
QY	241	SFSNSPQIIPRDRDIOCLIPCALDQDPYPMTDVAPRIGYBPKALLHSTFFPALQAO	300
Db	292	SFSNSPFIETFRDRDIOCLIPCALDQDPYFMTDVAIRIGHPRPALHSTFFPALQAO	351
QY	301	TKMSASDPNSSIIFLITPAKQIKTKVKNHAFSGGRDITIEHRQFGSNCVDVYSFWYLTFFL	360
Db	352	TKMSASDPNSSIIFLITPAKQIKSKVNNHAFSGGRDITIEHRQFGSNCVDVYSFWYLTFFL	411
QY	361	EDDDKLEQIRKDYTSYGAMLTGELKKALIEVLOPLIAHQAARKVETDIEYKEFWTPKRLS	420
Db	412	EDDRLEQIRKDYTSYGAMLTGELKKTLIDVLOPLIAHQAARKAVETIEYKEFWTPKRLS	471
QY	421	FDFO 424	
Db	472	FHQ 475	

RESULT 3			
09DBR9			
ID	09DBR9	PRELIMINARY;	PRT: 329 AA.
AC	09DBR9		
DT	01-JUN-2001 (TREMBLERel. 17, Created)		
DT	01-JUN-2001 (TREMBLERel. 17, Last sequence update)		
DT	01-MAR-2002 (TREMBLERel. 20, Last annotation update)		
DE	TrypToophanyl-tRNA synthetase.		
GN	WARS.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=PANCREAS;		
RC	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Aikawa T., Hara A., Fukunishi Y., Komio H., Adachi J., Fukuda S.,		
RA	Aikawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,		
RA	Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,		
RA	Schirral L.M., Staab F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gasteinich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima U., Mazzarelli U., Mombarts P.,
 RA Norioka P., Rung B., Ringwald C., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmshurst L.,
 RA Wymshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK007754; BAB25235.1; -
 DR MGD: MGI:104630; Mars.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR pfam: PF00579; tRNA-synt_1b.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 DR PROSITE: PS00178; AA TRNA LIGASE I; 1.
 SQ SEQUENCE 329 AA; 37613 MW; 701E702DC244C4A2 CRC64;

Query Match 72.5%; Score 1629; DB 11; Length 329;
 Best Local Similarity 92.1%; Pred. No. 2.4e-130;
 Matches 303; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

QY 96 MNQVLDAYENKRFYLYTGRGSSSEAMHGHLPFTKWLQDVFNPLVIOQTDEKYL 155
 DB 1 MNQVLDAYENKRFYLYTGRGSSSEAMHGHLPFTKWLQDVFNPLVIOQTDEKYL 60
 QY 156 MNDLTLDQAYGVDAVENAKDIIACGPDINKTFISDIDYMGMSGFKYKVVKIOKHTVFNQ 215
 DB 61 MNDLTLDQAYGVDAVENAKDIIACGPDINKTFISDIDYMGMSGFKYKVVKIOKHTVFNQ 120
 QY 216 VNGIFGFTSDSDICGKISPPAIOAPSPNSFPQIFRDRDIOCLIPCAIDODPYFMTD 275
 DB 121 VNGIFGFTSDSDICGKISPPAIOAPSPNSFPQIFRDRDIOCLIPCAIDODPYFMTD 180
 QY 276 VAPRIGYPPALHSHFFPALOGAOTKMSADPNSSIFLTDTKOIKTKVKNKAFSGGRD 335
 DB 181 VAPRIGYPPALHSHFFPALOGAOTKMSADPNSSIFLTDTKOIKTKVKNKAFSGGRD 240
 QY 336 TTEEHQFGGNCVDVDFMYLTFLEDDDLQEIIRKDYTGAMLTGELKKALIEVQPLI 395
 DB 241 TTEEHQFGGNCVDVDFMYLTFLEDDDLQEIIRKDYTGAMLTGELKKALIEVQPLI 300
 QY 396 AEHQARKKEVTDEIVKEFMTPRLSFDFQ 424
 DB 301 AEHQARKKEVTDEIVKEFMTPRLSFDFQ 329

RESULT 4
 ID 070184 PRELIMINARY; PRT; 305 AA.
 AC 070184;
 DT 01-ANG-1998 (T-EMBLrel. 07, Created)
 DT 01-ANG-1998 (T-EMBLrel. 07, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Trypophan-tRNA synthetase (Fragment).
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
 NCBI_TaxID=10141;
 RX SEQUENCE FROM N.A.
 RP STRAIN=HARTLEY; TISSUE=SPLEEN; Kobayashi Y.;
 RA Yang D., Goto R., Watanabe N.,
 RT "Identification and Cloning of Genes Whose Expressions are Elevated
 RT during DNCB-induced Guinea Pig Skin Delayed-type Hypersensitivity
 RT Reaction."
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB012222; BAA25288.1; -
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR002306; tRNA-synt_1b.

DR pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 KW Aminoacyl-tRNA synthetase.
 FT NON_TER
 SQ SEQUENCE 305 AA; 34838 MW; 0937164333780EB7 CRC64;

Query Match 68.4%; Score 1537; DB 11; Length 305;
 Best Local Similarity 94.4%; Pred. No. 1.4e-122;
 Matches 288; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 120 EAMHVGHLIPFTKWLQDVFNPLVIOQTDEKYLMDLTLDQAYGVDAVENAKDIIACG 179
 DB 1 EAMHVGHLIPFTKWLQDVFNPLVIOQTDEKYLMDLTLDQAYGVDAVENAKDIIACG 60
 QY 180 PDINKTFISDIDYMGMSGFKYKVVKIOKHTVFNQVIGIFRFTSDICGKISPPAIOA 239
 DB 61 PDINKTFISDIDYMGMSGFKYKVVKIOKHTVFNQVIGIFRFTSDICGKISPPAIOA 120
 QY 240 PPSFNSFPQIFRDRDIOCLIPCAIDODPYFMTDVAIRIGYPPALHSHFFPALOGA 299
 DB 121 PPSFNSFPQIFRDRDIOCLIPCAIDODPYFMTDVAIRIGYPPALHSHFFPALOGA 180
 QY 300 OTKMSADPNSSIFLTDKAKIKTKVKNKAFSGGRDTTIEHQFGGNCVDVDFMYLTF 359
 DB 181 OTKMSADPNSSIFLTDKAKIKTKVKNKAFSGGRDTTIEHQFGGNCVDVDFMYLTF 240
 QY 360 LEDDDLEQIRKDYTGAMLTGELKKALIEVQPLIAEHQARKKEVTDEIVKEFMTPR 419
 DB 241 LEDDDLEQIRKDYTGAMLTGELKKALIEVQPLIAEHQARKKEVTDEIVKEFMTPR 300
 QY 420 SFDFO 424
 DB 301 SFDFO 305

RESULT 5
 ID 0904Y1 PRELIMINARY; PRT; 430 AA.
 AC 0904Y1;
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE Trypophanyl-tRNA synthetase (AT21437p).
 GN AATS-TRP OR CG9735.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX SEQUENCE FROM N.A.
 RP MEDLINE=99250164; PubMed=10233165;
 RA Seshiah P., Andrew D.J.;
 RT "WRS-85D: A trypophanyl-tRNA synthetase expressed to high levels in
 RT the developing Drosophila salivary gland."
 RL Mol. Biol. Cell 10:1595-1608(1999).
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dreesen D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunco J., Paclob J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceiniker S.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF25156; AAF20166.1; -
 DR EMBL: AY075249; AAL68116.1; -
 DR FLYBase: FBgn0010803; Aats-trp.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_1b.
 DR pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.

DR TIGRFAMs; TIGR00233; ttps; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
 KW Aminoacyl-tRNA synthetase.
 SQ SEQUENCE 430 AA; 47985 MW; 2E3F9E9BC1E9979A CRC64;
 Query Match 61.5%; Score 1382; DB 5; Length 430;
 Best Local Similarity 61.6%; Pred. No. 3.6e-109;
 Matches 263; Conservative 63; Mismatches 85; Indels 16; Gaps 3;
 QY 10 DYKADCPGKPAFTSNHGPDA-----TEAE-----EDFVDEPTVQTSSAKGIDYD 54
 DB 3 DTKETVEGVEALTLNKPDAEPVETGTDAQAQEGATAPTEDVVDPMNVASSNDAGVDYD 62
 QY 55 KLIVRGSSKIDKELINIRATGQRPHHFLRGIFFSHRDMNOVLDAYENKKPFYLYTG 114
 DB 63 KLIRFSSSKIDKELINIRATGQRPHHFLRGIFFSHRDMNOVLDAYENKKPFYLYTG 122
 QY 115 RGSSEAMHVGHLIPFTFKMLDVFNVPLVLTQMTDEKYLMDLTDQAYGDAVENAKD 174
 DB 123 RGSSEAMHVGHLIPFTFKMLDVFNVPLVLTQMTDEKYLMDLTDQAYGDAVENAKD 182
 QY 175 IINCGPDINTFTFSDLDYGMSSGGFYKNVVKIOKAVTFNOVGIFGFTSDICGKISFP 234
 DB 183 IVAIGFVNKTFIFNNLEFVKGCPAMVONTIRIOKCVTFNOVGIFGFTSDICGKISFP 242
 QY 235 AIGAAPSFNSPQIPFRDRTDIOCLIPCAIDDPYFRMTDVAIRIGYPRKALHSTFP 294
 DB 243 AAGAAPSFNSPQIPFRDRTDIOCLIPCAIDDPYFRMTDVAIRIGYPRKALHSTFP 301
 QY 295 ALOGAQTKMSASDPNSSIFLTDTKAKOIKTKVKNKIASGSGDITIEHRFGNCDVVSFM 354
 DB 302 ALOGAQTKMSASDPNSSIFLTDTKAKOIKTKVKNKIASGSGDITIEHRFGNCDVVSFM 361
 QY 355 YLFFLEDDDKLEIQIRKDYTSGLMTSELKKALIEVLQPLAEHQARKKEVTDEIVKEM 414
 DB 362 LKFFLEDDDKLEIVRVAYSKGEMLTGEIKKLAVETLPIVEHQARKKITDEIVDKYF 421
 QY 415 TPKRLSF 421
 DB 422 ELRPLKF 428
 RESULT 6
 Q9U4Y0 PRELIMINARY; PRT; 420 AA.
 AC Q9U4Y0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Tyrophenyl-tRNA synthetase (Fragment).
 GN AATS--TRP OR CG9735.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX MEDLINE=99250164; PubMed=10233165;
 RA Seshiah P., Andrew D.J.;
 RT "WRS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in
 RT the developing Drosophila salivary gland.";
 RT Mol. Biol. Cell 10:1595-1608(1999).
 RL EMBL; AF125157; AAF20167.1; -;
 DR FlyBase; FBgn0010803; Aats-ctp.
 DR InterPro; IPR002305; CRNA-synt_1b.
 DR InterPro; IPR004142; CRNA-synt_1.
 DR InterPro; IPR002306; TTP CRNA-synt_1b.
 DR Pfam; PF00579; CRNA-synt_1b; 1.
 DR PRINTS; PR01039; TRNASNTTTRP.
 DR TIGRFAMs; TIGR00233; ttps; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
 KW Aminoacyl-tRNA synthetase.

PT NON TER. 1 1
 SQ SEQUENCE 420 AA; 46975 MW; 4AF7086AF426AB6 CRC64;
 Query Match 61.4%; Score 1379.5; DB 5; Length 420;
 Best Local Similarity 62.7%; Pred. No. 5.6e-109;
 Matches 257; Conservative 62; Mismatches 90; Indels 1; Gaps 1;
 QY 12 KADCPGKPAFTSNHGPDAFEAEDEFPVDPMTVQTSSAKGIDYDKLIVFGSSKIDKELIN 71
 DB 10 KPDAEPVETGTDAQAQEGATAPTEDVVDPMNVASSNDAGVDYDKLIRFSSSKIDKELIN 69
 QY 72 RIRATGQRPHHFLRGIFFSHRDMNOVLDAYENKKPFYLYTGSGSSEAMHVGHLIPFT 131
 DB 70 RPKIKGKPAHHIRRGVFMFSDHDLHTLTLRQKGFYLYTGSGSSEAMHVGHLIPFT 129
 QY 132 FTKMLDVFNVPLVLTQMTDEKYLMDLTDQAYGDAVENAKDINACGPDINTFTFSD 191
 DB 130 MTKMLDVFNVPLVLTQMTDEKYLMDLTDQAYGDAVENAKDINACGPDINTFTFSD 189
 QY 192 DYGMSSGGFYKNVVKIOKAVTFNOVGIFGFTSDICGKISFPALGAQTKMSASDPNSS 251
 DB 190 EFVKGCPAMVONTIRIOKCVTFNOVGIFGFTSDICGKISFPALGAQTKMSASDPNSS 249
 QY 252 DRDIDQCLIPCAIDDPYFRMTDVAIRIGYPRKALHSTFPALGAQTKMSASDPNSS 311
 DB 250 NR-KVHCLIPCAIDDPYFRMTDVAIRIGYPRKALHSTFPALGAQTKMSASDPNSS 308
 QY 312 IFETDPAKQIKTKVKNKIASGSGDITIEHRFGNCDVVSFMYLFFLEDDDKLEIQIRK 371
 DB 309 VYLLTDTPKQIKTKVKNKIASGSGDITIEHRFGNCDVVSFMYLFFLEDDDKLEIVR 368
 QY 372 DYSGAMLTGELKKALIEVLQPLAEHQARKKEVTDEIVKEMTPKRLSF 421
 DB 369 AYSKGEMLTGEIKKLAVETLPIVEHQARKKITDEIVDKYELRPLKF 418
 RESULT 7
 Q9VHG2 PRELIMINARY; PRT; 430 AA.
 AC Q9VHG2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Aats-ctp protein.
 GN AATS--TRP OR CG9735.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang C., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson R., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li Z., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., Meled M.P., Mepheson D.,
 RA Merkulov G., Mishina N.V., Mobarty C., Morris J., Moshell A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stimpson M., Strong R., Sun B.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003682; AG22136.1; -;
 DR FLYBase; FBgn0010803; Atcs-tpd.
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002306; ttp_cRNA-synt_1b.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PRO1039; TRNASYNTHTRP.
 DR TIGRFAMs; TIGR00233; trps; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
 SQ SEQUENCE 430 AA; 47971 MW; 2938EECC69B979F CRC64;

Query Match 61.4%; Score 1378; DB 5; Length 430;
 Best Local Similarity 61.4%; Pred. No. 7,8e-109;
 Matches 262; Conservative 64; Mismatches 85; Indels 16; Gaps 3;

10 DYKADPPGNPAFTSHGPD-----TEAE-----EDFVDPVTQTSAGKIDYD 54
 DB 3 DKKEIVVEGEALTLNGKPDABEVETGTDQAQEGATAPTEDVDVDPNNVASSNDAGVDYD 62
 QY 55 KLIIVRGSSKIDDELINRIERATGQRPHFLRGIFFSHDMNQVLDAYENKKEPFLYNG 114
 DB 63 KLIIVRGSSKIDDELINRIERATGQRPHFLRGIFFSHDMNQVLDAYENKKEPFLYNG 122
 QY 115 RGGSSAMVHGLIPRIFFTKMLQDVFNPLVIOPTDEKYLMDLTLDQAYGAVENAKD 174
 DB 123 RGGSSAMVHGLIPRIFFTKMLQDVFNPLVIOPTDEKYLMDLTLDQAYGAVENAKD 182
 QY 175 IIAAGDINKTIFSLDIDWMSGGFYKAVNKIQKHTENVQGIIFGTTSDICIGKISFP 234
 DB 183 IIAAGDINKTIFSLDIDWMSGGFYKAVNKIQKHTENVQGIIFGTTSDICIGKISFP 242
 QY 235 AIOAASFNSPQIRDRDIDLOCLPCALDODPYFRMRDVAIRIGYKPKALHSTFP 294
 DB 243 AIOAASFNSPQIRDRDIDLOCLPCALDODPYFRMRDVAIRIGYKPKALHSTFP 301
 QY 295 AIOAASFNSPQIRDRDIDLOCLPCALDODPYFRMRDVAIRIGYKPKALHSTFP 354
 DB 302 AIOAASFNSPQIRDRDIDLOCLPCALDODPYFRMRDVAIRIGYKPKALHSTFP 361
 QY 355 YLTFPEEDDKLEQIRKDYTGSMALTELGKALILEVLQPLIAHQARKVETDEIVEEM 414
 DB 362 YLTFPEEDDKLEQIRKDYTGSMALTELGKALILEVLQPLIAHQARKVETDEIVEEM 421
 QY 415 YLTFPEEDDKLEQIRKDYTGSMALTELGKALILEVLQPLIAHQARKVETDEIVEEM 421
 DB 422 YLTFPEEDDKLEQIRKDYTGSMALTELGKALILEVLQPLIAHQARKVETDEIVEEM 421

RESULT 8
 O9SR15 PRELIMINARY; PRT; 402 AA.
 AC O9SR15, 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative tryptophanyl-tRNA synthetase.
 GN F018.7 OR AF3604600.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosida II; Brassicales; Brassicaceae; Arabidopsie.
 OX NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
 RA Roming C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Niernan W.C., Frazer C.M.,
 RT Arabidopsis thaliana chromosome III BAC F7018 genomic sequence.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tortum M., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai U., Kim C., Lam B., Lin J.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.D., Sakurai T., Satou M.,
 RA Seki M., Shim P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT Arabidopsis full length cDNA clones.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC011437; AA04890.1; -;
 DR EMBL; AY080709; AA085027.1; -;
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002306; ttp_cRNA-synt_1b.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PRO1039; TRNASYNTHTRP.
 DR TIGRFAMs; TIGR00233; trps; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
 KW Aminoacyl-tRNA synthetase.
 SQ SEQUENCE 402 AA; 45754 MW; B9BEA75E5D6CD15 CRC64;

Query Match 57.9%; Score 1301; DB 10; Length 402;
 Best Local Similarity 60.3%; Pred. No. 2.5e-102;
 Matches 240; Conservative 66; Mismatches 88; Indels 4; Gaps 2;

29 DATAE--EDFDPMTVQTSAGKIDYDLYRFGSSKIDKLINRIERATGQRPHFLR 86
 DB 7 DEREAASSQVNVNPEVSAKDGKIDYDLYRFGSSKIDKLINRIERATGQRPHFLR 66
 QY 87 RGIFFSHRDMNOVLDAVENKKEPFLYLTGSGPSEAMVHGLIPRIFFTKMLQDVFNPLVI 146
 DB 67 RGIFFSHRDMNOVLDAVENKKEPFLYLTGSGPSEAMVHGLIPRIFFTKMLQDVFNPLVI 126
 QY 147 QMTDEKYLMDLTLDQAYGAVENAKDIIAAGDINKTIFSLDIDWMSGGFYKAVNK 206
 DB 127 QMTDEKYLMDLTLDQAYGAVENAKDIIAAGDINKTIFSLDIDWMSGGFYKAVNK 184
 QY 207 ICKHTENVQGIIFGTTSDICIGKISFPALQABSPNSFPQIRDRDIDLOCLCAIDQ 266
 DB 185 VIKCVTLNKAQKIFGSGEDPFLAKSPFPVQVNPSPSPFPFLPGKDLRLICPAIDQ 244
 QY 267 DVFPMTRDVAIRIGYKPKALHSTFPALQAGTKMSASDNNISIFLDTAKQIKTKVN 326
 DB 245 DVFPMTRDVAIRIGYKPKALHSTFPALQAGTKMSASDNNISIFLDTAKQIKTKVN 304
 QY 327 KIAFSGGRDTIEHRQFGKNCVDVSPMYLTFLEDDDKLEQIRKDYTGSMALTELGKKA 386
 DB 305 RYAFSGGQDSIEKHRELGNLELVDPVKYLSFLEDDDKLEQIRKDYTGSMALTELGKKA 364
 QY 387 LIEVLQPLIAHQARKVETDEIVEEMTPRLTSPDQ 424
 DB 365 LIEVLQPLIAHQARKVETDEIVEEMTPRLTSPDQ 402

RESULT 9

```

Q9UR2      PRELIMINARY;      PRT;      324 AA.
ID Q9UR2
AC Q9UR2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Y80D3A.1 protein.
GN Y80D3A.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B.R.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=9069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating Biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AL132853; CAB60439.1; -.
DR InterPro; IPR002305; CRNA-synt_1b.
DR InterPro; IPR002306; TTP-CRMA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PRO1039; TRNASyntTRP.
SQ SEQUENCE 324 AA; 36289 MW; 6E687E6E420EC12 CRC64;

Query Match      45.7%; Score 1027; DB 5; Length 324;
Best Local Similarity 64.9%; Pred. No. 3,8e-79;
Matches 194; Conservative 39; Mismatches 64; Indels 2; Gaps 1;

QY 27 GPDATAEEDPVPWTVQTSASAKIDYDKLIVFGSSKIDKELINRTERATGQRPHFLR 86
DB 23 GGGQVEDEEDRVTPEWETTKATGIDYDKLIVFGCKRLDEILARFERYTGKASPMR 82
QY 87 RGFPSHRDNNQVDAENKKPFILYGRGSSSAMVGHILPIFKMLQDVNVNVLVI 146
DB 83 RGMFAIRDLTALIDREOGPFYLYGRGSSGSLHGVPIFKMLQEVFVPLVI 142
QY 147 QMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFIDNKTFFISLDLYMGSSGFYKAVK 206
DB 143 QMTDDEKFLMKDKVDAKKAARENMDIISVGFDPKTFIENNFDI--WCPPEYENIVK 200
QY 207 IQKXVTFNOKGIFGFTDSDCIKISFPALQAPSFNSFPQIFRRTDIQCLIPCAIDQ 266
DB 201 IMKVNTNQAARAIFFGTFPEDCIGAKAPPAVEAACPFSFQIFGKENDIPCLIPCAIDQ 260
QY 267 DPFYRMTDVAAPRIGYKPKPALHSTFPFPAQAGQTKASDPNSIFLTTAKQIKTKV 325
DB 261 DPFYRMTDVAAPRIGYKPKPALHSTFPFPAQAGQTKASDPNSIFLTTAKQIKTKV 319

RESULT 10
Q8S0Y5      PRELIMINARY;      PRT;      385 AA.
ID Q8S0Y5
AC Q8S0Y5;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Tryptophanyl tRNA synthetase.
GN EC011_0530.
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxId=6035;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA Prenster G., Barbe V., Peyrethallade E., Broctier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivasres C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi.";
RL Nature 414:450-453(2001).
DR EMBL; AL590450; CAD25963.1; -.
SQ SEQUENCE 385 AA; 44188 MW; 2865C3AC95FC859 CRC64;

Query Match      41.9%; Score 942; DB 5; Length 385;
Best Local Similarity 46.1%; Pred. No. 8.3e-72;
Matches 179; Conservative 76; Mismatches 123; Indels 10; Gaps 3;

QY 35 EDFVDPWTVQTSASAK----GIDYDKLIVFGSSKIDKELINRTERATGQRPHFLRGIF 90
DB 3 EQRTFMDVEVSTDEVPVAIDYDKLINQGCCKFNQALADRLKLSGKPAHYFFRGIV 62
QY 91 FSHRDNNQVDAENKKPFILYGRGSSSAMVGHILPIFKMLQDVNVNVLVIQMTD 150
DB 63 FAHRDPNLLDEIANNRPFLYLYGRGSSSKTWHIGHTIPFLCKYQDAKRLVLIQTD 122
QY 151 DEKYLWKDLTLDOAYGDAVENAKDIIACGFIDNKTFFISLDLYMGSSGFYKAVVYIOKH 210
DB 123 DEKFLMSKMLEDAAMAYGRNIDYALGFDPLKLTIFSNVE--ASHHFEENILKISTK 179
QY 211 VTFNOKGIFGFTDSDCIKISFPALQAPSFNSFPQIFRRTDIQCLIPCAIDDPYF 270
DB 180 ININEMAIKVYGFPMSSNIGQVGPFAKEIAPCFSSSFRFGK---GAMCLVPAVDDPFR 236
QY 271 RMTRDVAAPRIGYKPKPALHSTFPFPAQAGQTKASADPNSSIFLTTAKQIKTKVKNAR 330
DB 237 RLARDAKAKLGEKKPSIYVSLPDLKGVKRRKASADPNSSIYLDQAQDTIRKIIAYAY 296
QY 331 SGGRTDIEEHROGNCVDVVSFMYLTFPLEDDDKLEQRIKDYTSGLMLTGLKALIEV 390
DB 297 SGGRTDIEEHREKGGIDVDVPEYLYKFLDDQELKRSGLIKELITSKEMKCVV 356
QY 391 LQPLIAHQARRKEVTDEIVEEMTPRK 418
DB 357 IQEFSRYQESRKRYTDDDLARFIDINK 384

RESULT 11
Q8U453      PRELIMINARY;      PRT;      385 AA.
ID Q8U453
AC Q8U453;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Tryptophanyl tRNA synthetase.
GN PF0241.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxId=2261;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010149; AL480365.1; -.
KW Aminocacyl-tRNA synthetase; Complete proteome.
SQ SEQUENCE 385 AA; 45178 MW; 3A7A628958200CC CRC64;

Query Match      36.8%; Score 826.5; DB 17; Length 385;
Best Local Similarity 45.9%; Pred. No. 5.6e-62;
Matches 180; Conservative 69; Mismatches 116; Indels 27; Gaps 10;

```

```

QY 34 BEDF-VDPMWVOTSSAKGIDYDLIVRFGSSKIDKELINRERATQGRPHHFLRGIFPS 92
DB 2 EEEFKTPEVEBEI-----IDYNKLIQGFISPLTDLERTARLTSELPFRFRKFFPS 57
QY 93 HEDMNOVLDAYENKPEFYLYTGRGSSSEAMHVGHLIPFTFKWLDQVFNVLQMTDDE 152
DB 58 HBDYVDVLDDYQKGFPLYTGSGPSG-PWHIGHIIPFPATKWLQKCFDNLVLIQTDD 116
QY 153 KTLMD-LTLDOAYGAVENAKDIIACGPDINTFTFSDLDYMGSSGYKAVKIQKIV 211
DB 117 KFLFENLFFEDTKWYAONILDIIVAGDPDPTFFLQNSF-----TKLYEAMAPIAKKI 172
QY 212 TENOVGIFGFTDSDICIGISPAPAOAPSFSNFPQIFRDRDIOCLIPCAIDODPYR 271
DB 173 NNSMAKAVGFTGOSKIGMIFPAIOAPTF-----FEK-----RCLIPALIDODPYR 222
QY 272 MTRDVAIRIGYKPPALHSTFPFPAIOQAOTKMSASDPNSSIFLTDTAKOIKTKVNHAFS 331
DB 223 LORDFAESLIGYVYTAALHSHKFPPLTGLBGKMSASKEPTAIYLTDPBEAGKKIMWFALT 282
QY 332 GGRDTIEHRQRCNGCDVVSFMYLTFPLEDDKLEQIRKDY---TSGMLTGLKAL 388
DB 283 GGQPTLKEREKRGKNEKCVFEMLEIFPEPDK--KLMEYYACKNGELGCECKRYLI 340
QY 389 EVLQPLIAHQAARKKEVTDEIVKEFMTPRKLS 420
DB 341 QRVQEFLEKHQEKRK-AEKLVKFKYTKLA 371

```

RESULT 12

```

ID 09U1F5 PRELIMINARY; PRT; 490 AA.
AC 09U1F5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Tryptophanyl-CRNA synthetase.
GN L1063.04.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=FRIDLIN;
RA Bothe G., Pohl T., Ivens A.C., Lawson D., Quail M., Rajandream M.A.,
RA Barrell B.G.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=FRIDLIN;
RX MEDLINE=98146435; PubMed=9477341;
RX Ivens A.C., Lewis S.M., Bagnerzdeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RA "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; AL121862; CAB58393.1; -.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; tRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b.
DR PRINTS; PRO1039; TRNASYNTHTRP.
DR PROSITE; PS00178; AA_TRNA_LIGASE_1.
KW Aminoacyl-tRNA synthetase.
SQ SEQUENCE 490 AA; 5490 MW; 172C68622C511D3D CRC64;

```

Query Match 35.6%; Score 800.5; DB 5; Length 490;
 Best Local Similarity 37.7%; Pred. No. 1.3e-59;
 Matches 174; Conservative 89; Mismatches 133; Indels 65; Gaps 11;

```

QY 19 NPAFTSNHGPDAFEEDPDEPTVOTSSAKGIDYDLIVRFGSSKID-----KELIN 71
DB 33 DPSEPVQHPRDDGAGAEVITPPVVAKAGPQGINRYVLTIFAEKRMDDGARGHMDVNA 92

```

```

QY 72 RIERAT-----GGRP-----HFLRGCI 89
DB 93 KCRKKTMTTSPSANAAGVATPVLDVAPGADALQAASPAHQOQRPALALHFFHRDI 152
QY 90 FSHEDMNOVLDAYE-----NKKPFYLYTGRGSSSEAMHVGHLIPFTFKWLDQVFNVL 144
DB 153 ARSHRDLHKAALVDIEASVKTGRSVFLYTGROPSAGTMHLGHVLPMLTKVLDVFSPL 212
QY 145 VLOMTDEKYLKDLTDOAYGDV--ENAKIITACGPDINTFTFSDLDYMGSSGYFK 202
DB 213 VQITDDEKFLRDPDPEBKADELIRSNIKDIIENFMRPTFLFRNHYG--DWP 269
QY 203 NVKTIQKHTVFNQVKGIFGFTDSDICIGISPAPAOAPSFSNFPQIFRDRDIOCL 260
DB 270 TVLRQSRMTGNAVHTGLTITSDVNGKLAFPATQAAPCFSTAFKRVLRNGDR-PKRCII 328
QY 261 PCAIDODPYRMTRDVAIRIGYKPPALHSTFPFPAIOQAOTKMSAS-DPNSSIFLTDTAK 319
DB 329 PCAIDODPEFVLTRAALRLKQLPALHSTFKLPALKGLHMGSSSAEKGVITLHDTDK 388
QY 320 QIKTKVNHAFSGGKADTIEHRQRCNGCDVVSFMYLTFPLEDDKLEQIRKDYTSGANL 379
DB 389 QVRKRL-RRAFSGGATLEQMOGTANLELDVAYOYLRFECBDPTLPAVOTGRYRSGTLN 447
QY 380 TSELKALTE-VLOPLIAHQAARKKEVTDEIVKEFMTPRKLS 419
DB 448 SEVVDLMAADCIREVLDHMRERRATVDDDVYEFGRIRDI 488

```

RESULT 13

```

ID 095295 PRELIMINARY; PRT; 136 AA.
AC 095295;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Tryptophanyl-CRNA synthase (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RA Wintneroe A.K., Fredholm M., Davies W.;
RT Evaluation and characterization of a porcine small intestine cDNA
RT library.
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z81267; CAB03585.1; -.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR002306; tRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b.
DR PRINTS; PRO1039; TRNASYNTHTRP.
FT NON TER 1 1
FT NON TER 136 136
SQ SEQUENCE 136 AA; 15404 MW; 951F75D1B9CD0617 CRC64;

```

Query Match 30.4%; Score 682; DB 6; Length 136;
 Best Local Similarity 94.1%; Pred. No. 2.5e-50;
 Matches 128; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

QY 234 PALQAAPSFSNFPQIFRDRDIOCLIPCAIDODPYRMTRDVAIRIGYKPPALHSTFP 293
DB 1 PALQAAPSFSNFPQIFRDRDIOCLIPCAIDODPYRMTRDVAIRIGYKPPALHSTFP 60
QY 294 PALQAOTKMSASDPNSSIFLTDTAKOIKTKVNHAFSGGRTIEHRQRCNGCDVVSF 353
DB 61 PALQAOTKMSAXAPNSSIFLTDTAKOIKTKVNHAFSGGRTIEHRQRCNGCDVVSF 120
QY 354 MYLTFLEDDKLEQIRK 369
DB 121 MYLTFLEDDKLEQIRK 136

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:27:20 ; Search time 27.2194 Seconds

(without alignments)
3035.516 Million cell updates/sec

Title: US-09-813-718-10_COPY_71_471

Perfect score: 2116
Sequence: 1 SNHGDATAEDFDVDPWTV.....VTDEIKERMTPRKLSFDQ 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1957	92.5	475	11 Q9DC65	Q9DC65 mus musculus
2	1957	92.5	481	11 Q99J58	Q99J58 mus musculus
3	1629	77.0	329	11 Q9DBR9	Q9DBR9 mus musculus
4	1537	72.6	305	11 Q70184	Q70184 cavia porce
5	1374.5	65.0	420	5 Q9UAY0	Q9UAY0 drosophila
6	1374.5	65.0	430	5 Q9UAY1	Q9UAY1 drosophila
7	1370.5	64.8	430	5 Q9VHG2	Q9VHG2 drosophila
8	1301	61.5	402	10 Q9SR15	Q9SR15 arabidopsis
9	1027	48.5	324	5 Q9UR12	Q9UR12 caenorhabditis
10	942	39.5	385	5 Q8SOY5	Q8SOY5 encephalito
11	826.5	34.1	385	17 Q8U453	Q8U453 pyrococcus
12	795.5	37.6	490	5 Q9U1F5	Q9U1F5 leishmania
13	682	32.2	136	6 Q9S295	Q9S295 sus scrofa
14	453	21.4	17	6 Q8TYF7	Q8TYF7 mechanopyru
15	329.5	15.6	437	17 Q8TUA1	Q8TUA1 mechanosarc
16	310	14.7	111	5 Q95YL8	Q95YL8 encephalito

17	296.5	14.0	136	6 Q9TS88	Q9TS88 bos taurus
18	294	13.9	157	5 Q9US33	Q9US33 trypanosoma
19	286	13.5	109	4 Q9UD15	Q9UD15 homo sapien
20	269.5	12.7	513	17 Q9HN83	Q9HN83 halobacteri
21	191	9.0	324	17 Q8TX22	Q8TX22 mechanopyru
22	188	8.9	364	17 Q9YA64	Q9YA64 aeropyrum p
23	181.5	8.6	331	17 Q979Z1	Q979Z1 thermoplasm
24	171.5	8.1	351	17 P93018	P93018 arabidopsis
25	171.5	8.1	385	10 Q8S9J2	Q8S9J2 arabidopsis
26	167	7.9	351	16 Q9RVDS	Q9RVDS deinococcus
27	164.5	7.8	895	10 Q9SGN2	Q9SGN2 arabidopsis
28	158.5	7.5	375	17 Q9V027	Q9V027 pyrococcus
29	156	7.4	372	17 Q8ZY77	Q8ZY77 pyrobaculum
30	154.5	7.3	102	10 Q07119	Q07119 halobacteri
31	153.5	7.3	317	17 Q8TS11	Q8TS11 mechanosarc
32	152.5	7.2	682	5 Q9N9B8	Q9N9B8 leishmania
33	149.5	7.1	327	17 Q9HN62	Q9HN62 halobacteri
34	149.5	7.1	408	10 P93363	P93363 nicotiana t
35	148	7.0	316	17 Q8ZW77	Q8ZW77 pyrobaculum
36	146	6.9	375	17 Q58739	Q58739 pyrococcus
37	145	6.9	332	17 Q9HK73	Q9HK73 thermoplasm
38	145	6.9	337	16 Q8Z0Y0	Q8Z0Y0 salimonia
39	142.5	6.7	375	17 Q8U2H3	Q8U2H3 pyrococcus
40	142.5	6.7	528	11 Q91WQ3	Q91WQ3 mus musculus
41	140	6.6	337	16 Q8ZK00	Q8ZK00 salimonia
42	139	6.6	294	17 Q96YV3	Q96YV3 sulfolobus
43	135	6.4	365	16 Q9KRV7	Q9KRV7 vibrio chol
44	134.5	6.4	525	5 Q9VAV60	Q9VAV60 drosophila
45	130.5	6.2	419	16 Q92BB1	Q92BB1 listeria in

ALIGNMENTS

RESULT 1
ID Q9DC65 PRELIMINARY; PRT; 475 AA.

AC Q9DC65;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DE Adult male lung cDNA, RIKEN full-length enriched library,
clone:1200002C07, full insert sequence.
GN WARS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LUNG; MEDLINE=21085660; PubMed=11217851;
RX Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakita I.,
Saio T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saio R.,
Kadoya K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohzuki S.,
Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL, AK004541, BAB23357.1, -.

DR MGD; MGI.104630; MARS.
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002306; tTP_CRNA-synt_1b.
 DR InterPro; IPR000738; WHEP-TRS.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR Pfam; PF00458; WHEP-TRS; 1.
 DR PRINTS; PRO1039; TRNASYNTHTRP.
 DR TIGRFAMs; TIGR00233; ttps; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
 DR PROSITE; PS00762; WHEP-TRS; 1.
 DR SQUENCE 475 AA; 53641 MW; C3467FE85521DE4C CRC64;

Query Match 92.5%; Score 1957; DB 11; Length 475;
 Best Local Similarity 91.5%; Pred. No. 3.6e-159;
 Matches 366; Conservative 20; Mismatches 14; Indels 0; Gaps 0;

QY 2 NHGPDATAEEDFVDPMTVQTSKAGIDYDKLIVRFSSKIDKELINRIERATGQRPHP 61
 DB 76 NCSDATKASEDFVDPMTVQTSKAGIDYDKLIVQFSSKIDKELINRIERATGQRPHP 135
 QY 62 LRKGIFFSHRDMQVLDAYENKKPFYLYTGRGSSSAMVGHILPIFTKMLQDVNVPL 121
 DB 136 LRKGIFFSHRDMQVLDAYENKKPFYLYTGRGSSSAMVGHILPIFTKMLQDVNVPL 195
 QY 122 VIQMTDDEKYLMDLTLDOAYGDVAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKNV 181
 DB 196 VIQMSDEKYLMDLTLDOAYSTVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKNV 255
 QY 182 VKIQKHTFNOVGIFGFTSDICIGKISFPALQAPAFSNSFPQIFRDRDIOCLIPCAI 241
 DB 256 VKIQKHTFNOVGIFGFTSDICIGKISFPALQAPAFSNSFPQIFRDRDIOCLIPCAI 315
 QY 242 DDDPYFRMTDVAPRIGYPRPALHSTFPALQAGATKMSASDPNSIFLTDTAKQIKTK 301
 DB 316 DDDPYFRMTDVAPRIGYPRPALHSTFPALQAGATKMSASDPNSIFLTDTAKQIKTK 375
 QY 302 VNHGAFSGGSDTIEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYSGMLTGEELK 361
 DB 376 VNHGAFSGGSDTIEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYSGMLTGEELK 435
 QY 362 KALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401
 DB 436 KALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 475

RESULT 2

Q99J58 PRELIMINARY; PRT; 481 AA.
 AC 099J58;
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE Trypophanyl-tRNA synthetase.
 GN MARS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MAMMARY TUMOR. WAP-TAG MODEL. 5 MONTHS OLD;
 RA Straussberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC003450; AA03450.1; -.
 DR MGD; MGI.104630; MARS.
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002306; tTP_CRNA-synt_1b.
 DR InterPro; IPR000738; WHEP-TRS.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR Pfam; PF00458; WHEP-TRS; 1.
 DR PRINTS; PRO1039; TRNASYNTHTRP.

DR TIGRFAMs; TIGR00233; ttps; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
 DR PROSITE; PS00762; WHEP-TRS; 1.
 DR Aminoacyl-tRNA synthetase.
 DR SQUENCE 481 AA; 54325 MW; A754E1DDF58E2EF3 CRC64;

Query Match 92.5%; Score 1957; DB 11; Length 481;
 Best Local Similarity 91.5%; Pred. No. 3.7e-159;
 Matches 366; Conservative 20; Mismatches 14; Indels 0; Gaps 0;

QY 2 NHGPDATAEEDFVDPMTVQTSKAGIDYDKLIVRFSSKIDKELINRIERATGQRPHP 61
 DB 76 NCSDATKASEDFVDPMTVQTSKAGIDYDKLIVQFSSKIDKELINRIERATGQRPHP 135
 QY 62 LRKGIFFSHRDMQVLDAYENKKPFYLYTGRGSSSAMVGHILPIFTKMLQDVNVPL 121
 DB 136 LRKGIFFSHRDMQVLDAYENKKPFYLYTGRGSSSAMVGHILPIFTKMLQDVNVPL 195
 QY 122 VIQMTDDEKYLMDLTLDOAYGDVAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKNV 181
 DB 196 VIQMSDEKYLMDLTLDOAYSTVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKNV 255
 QY 182 VKIQKHTFNOVGIFGFTSDICIGKISFPALQAPAFSNSFPQIFRDRDIOCLIPCAI 241
 DB 256 VKIQKHTFNOVGIFGFTSDICIGKISFPALQAPAFSNSFPQIFRDRDIOCLIPCAI 315
 QY 242 DDDPYFRMTDVAPRIGYPRPALHSTFPALQAGATKMSASDPNSIFLTDTAKQIKTK 301
 DB 316 DDDPYFRMTDVAPRIGYPRPALHSTFPALQAGATKMSASDPNSIFLTDTAKQIKTK 375
 QY 302 VNHGAFSGGSDTIEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYSGMLTGEELK 361
 DB 376 VNHGAFSGGSDTIEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYSGMLTGEELK 435
 QY 362 KALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401
 DB 436 KALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 475

RESULT 3

Q9DBR9 PRELIMINARY; PRT; 329 AA.
 AC Q9DBR9;
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
 DE Trypophanyl-tRNA synthetase.
 GN MARS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai U.; Shingawa A.; Shibata K.; Yoshino M.; Itoh M.; Ishii Y.;
 RA Atakawa T.; Hara A.; Fukunishi Y.; Kono H.; Adachi S.; Fukuda S.;
 RA Aizawa K.; Izawa M.; Nishi K.; Kiyosawa H.; Kondo J.; Yamahata I.;
 RA Saito T.; Okazaki Y.; Gojobori T.; Bono H.; Kasukawa T.; Saito R.;
 RA Kadota K.; Matsuda H.A.; Ashburner M.; Batalov S.; Casavant T.;
 RA Fleischnann W.; Gaasterland T.; Glass C.; King B.; Kochiwa H.;
 RA Kuehl P.; Lewis S.; Matsuo Y.; Nikaido I.; Pesole G.; Quackenbush J.;
 RA Schriml L.M.; Staudli F.; Suzuki R.; Tomita M.; Wagner L.; Washio T.;
 RA Sakai K.; Okido T.; Furuno M.; Aono H.; Baldarelli R.; Barsch G.;
 RA Blake J.; Boffelli D.; Bojunga N.; Carninci P.; de Bonaldo M.F.;
 RA Brownstein M.J.; Bult C.; Fletcher C.; Fujita M.; Gariboldi M.;
 RA Gustincich S.; Hill D.; Hofmann M.; Hume D.A.; Kamiya M.; Lee N.H.;
 RA Lyons P.; Marchionni L.; Mashima U.; Mazarrelli J.; Mombaerts P.;
 RA Nordone P.; Ring B.; Ringwald M.; Rodriguez I.; Sakamoto K.-F.;
 RA Sasaki H.; Sato K.; Schonbach C.; Seya T.; Shibata Y.; Storch K.-F.;
 RA Suzuki H.; Toyooka K.; Wang K.H.; Weitz C.; Whitaker C.; Wilming L.;
 RA Wyshaw-Boris A.; Yoshida K.; Hasegawa Y.; Kawai H.; Kohsaki S.;

RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK007754; BAB25235.1; -.
 DR MGI; MGI:104630; Wats.
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002306; tRNA-synt_1b.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PRO1039; TRNASYNTHTRP.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
 SQ SEQUENCE 329 AA; 37613 MW; 701E702DC244CA2 CRC64;

Query Match 77.0%; Score 1629; DB 11; Length 329;
 Best Local Similarity 92.1%; Pred. No. 2.7e-131;
 Matches 303; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

QY 73 MNQVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFIETKMLQDVENVPLVIQMTDEKYL 132
 DB 1 MNQILDAYENKKPFYLYTGRGSSSEAMHVGHLIPFIETKMLQDVENVPLVIQMSDEKYL 60
 QY 133 WKDLTLDQAYGDAVENAKOIIACGPDINTFTFSDLDYMGSSGFYKNVVKIÖKHVTENQ 192
 DB 61 WKDLTLEQAYSYTVENAKOIIACGPDINTFTFSDLDYMGSSGFYKNVVKIÖKHVTENQ 120
 QY 193 VKGIFGFTSDICIGKISFPALQAPSFNSPQIFPRDRDIOCLIPCAIDODPYFRMTD 252
 DB 121 VKGIFGFTSDICIGKISFPALQAPSFNSPQIFPRDRDIOCLIPCAIDODPYFRMTD 180
 QY 253 VAPRIGYPRPALHSTFFPALQAGOTKMSASDPNSSIFLTPDAKOIKTKVNHGASGGSD 312
 DB 181 VAPRIGHPRPALHSTFFPALQAGOTKMSASDPNSSIFLTPDAKOIKSVNHGASGGSD 240
 QY 313 TIEEHROFGSCDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMTELKKALILEVLOPLI 372
 DB 241 TIEEHROFGSCDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMTELKKALILEVLOPLI 300
 QY 373 AEHQARRKEVTDEIYKEFTPRKLSFDFQ 401
 DB 301 AEHQARRKAIVTEIYKEFTPRKLSFDFQ 329

RESULT 4

QY 070184 PRELIMINARY; PRT; 305 AA.
 AC 070184;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Trypophan-tRNA synthetase (Fragment).
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HARTLEY; TISSUE=SPLEEN;
 RA Yang D., Goto K., Watanabe N., Kobayashi Y.;
 RT "Identification and Cloning of Genes Whose Expressions are Elevated
 RT during DMCB-induced Guinea Pig Skin Delayed-type Hypersensitivity
 RT Reaction."
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBD databases.
 DR EMBL; AB012222; BAA25288.1; -.
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR002306; tRNA-synt_1b.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PRO1039; TRNASYNTHTRP.
 KW Aminoacyl-tRNA synthetase.
 FT NON TER 1
 SQ SEQUENCE 305 AA; 34838 MW; 0937164333780EB7 CRC64;

Query Match 72.6%; Score 1537; DB 11; Length 305;
 Best Local Similarity 94.4%; Pred. No. 1.8e-123;

Matches 288; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 97 EAMHVGHLIPFIETKMLQDVENVPLVIQMTDEKYLMDLTDQAYGDAVENAKOIIACG 156
 DB 1 EAMHVGHLIPFIETKMLQDVENVPLVIQMSDEKYLMDLTDQAYGDAVENAKOIIACG 60
 QY 157 PINTNFIETSDLDYMGSSGFYKNVVKIÖKHVTENQVIGIFGFTSDICIGKISFPALQAA 216
 DB 61 PINTNFIETSDLDYMGSSGFYKNVVKIÖKHVTENQVIGIFGFTSDICIGKISFPALQAA 120
 QY 217 PSFNSPQIFPRDRDIOCLIPCAIDODPYFRMTDVAIPRIGYPRPALHSTFFPALQGA 276
 DB 121 PSFNSPQIFPRDRDIOCLIPCAIDODPYFRMTDVAIPRIGYPRPALHSTFFPALQGA 180
 QY 277 ÖTKMSASDPNSSIFLTPDAKOIKTKVNHGASGGSDTIEEHROFGSCDVDSFMYLTFF 336
 DB 181 ÖTKMSASDPNSSIFLTPDAKOIKTKVNHGASGGSDTIEEHROFGSCDVDSFMYLTFF 240
 QY 337 LEDDDKLEQIRKDYTSGLMTELKKALILEVLOPLIAEHQARRKEVTDEIYKEFTPRKL 396
 DB 241 LEDDDKLEQIRKDYTSGLMTELKKALILEVLOPLIAEHQARRKEVTDEIYKEFTPRKL 300
 QY 397 SFDQ 401
 DB 301 SFDQ 305

RESULT 5

QY 09U4Y0 PRELIMINARY; PRT; 420 AA.
 AC 09U4Y0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Trypophanyl-tRNA synthetase (Fragment).
 GN AATS-TRP OR CG9735.
 GN Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta.
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99250164; PubMed=10233165;
 RA Seshiah P., Andrew D.O.;
 RT "WRS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in
 RT the developing Drosophila salivary gland."
 RL Mol. Biol. Cell 10:1595-1608(1999).
 DR EMBL; AF125157; AAF20167.1; -.
 DR FlyBase; FBgn0010803; Aats-trp.
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002306; tRNA-synt_1b.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PRO1039; TRNASYNTHTRP.
 DR TIGRfam; TIGR00233; trps; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
 KW Aminoacyl-tRNA synthetase.
 FT NON TER 1
 SQ SEQUENCE 420 AA; 46975 MW; 4AF7088AF426A86 CRC64;

Query Match 65.0%; Score 1374.5; DB 5; Length 420;
 Best Local Similarity 64.8%; Pred. No. 2.4e-109;
 Matches 254; Conservative 61; Mismatches 76; Indels 1; Gaps 1;

QY 7 ATEAEEDVDPMTYQTSAGKIDYDKLIVRGSSKIDELINRTERATGQRPHHFLRRGI 66
 DB 28 ATEAEEDVDPMTYQTSAGKIDYDKLIVRGSSKIDELINRTERATGQRPHHFLRRGI 87
 QY 67 FFSHDMNQVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFIETKMLQDVENVPLVIQMT 126
 DB 88 FFSHDLTILTLRQKPFYLYTGRGSSSEAMHVGHLIPFIETKMLQDVENVPLVIQMT 147

QY 127 DDEKYLKMDLTLDQAVGDAVENAKDIACGPDINKFIFSDLDYMGMSGGFYKNNVYIOK 186
DB 148 DDEKYLKMDLTLDQAVGDAVENAKDIACGPDINKFIFSDLDYMGMSGGFYKNNVYIOK 207
QY 187 HTFENOVKGIFFGTSDSDICISFPALQOAPSPFSNPFQIFRDRDIOCLIPALDODPY 246
DB 208 CVTFENOVKGIFFGTSDSDICISFPALQOAPSPFSNPFQIFRDRDIOCLIPALDODPY 266
QY 247 FFRTRDVAAPRIGYIPKALHSTFFPALQOAKTKMSADPNSSIFLDTAOKITKNNKA 306
DB 267 FFRTRDVAAPRIGYIPKALHSTFFPALQOAKTKMSADPNSSIFLDTAOKITKNNKA 326
QY 307 FSGGRDTEHROFGNCVDVSMYLTFFLEDDDKLEOIRKDYTSGAMLTGELKALIE 366
DB 327 FSGGRVTVVEHRLKGGVPEVDVSYOLKFFLEDDDALEVRVAVSKGEMLTGELKALIE 386
QY 367 VLOPLIAEHQARKEVTDEIVKEFMPRKLSP 398
DB 387 TLTPIVEHQARKLITDEVLDKYEFLRPLKF 418
RESULT 6
Q904Y1 PRELIMINARY; PRT; 430 AA.
ID Q904Y1
AC Q904Y1
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Trypophanyl-tRNA synthetase (At21437p).
GN AATS-TRP OR CG9735.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99250164; PubMed=10233165;
RA Seshiah P., Andrew D.J.,
RT "WRS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in
the developing Drosophila salivary gland."
RL Mol. Biol. Cell 10:1595-1608(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarni H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Munoz J., Paclebo J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF125156; AAF20166.1; -;
DR EMBL, AY075249; AAL68116.1; -;
DR FLYBase; FBgn0010803; Aats-trp.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; Trp tRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PRO1039; TRNASYNTRP.
DR TIGRFA; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA tRNA LIGASE_1; 1.
KW Aminoacyl-tRNA synthetase.
SO SEQUENCE 430 AA; 47985 MW; 2E3F9E9BC1E9979A CRC64;
Query Match 65.0%; Score 1374.5; DB 5; Length 430;
Best Local Similarity 64.8%; Pred. No. 2.5e-109;
Matches 254; Conservative 61; Mismatches 76; Indels 1; Gaps 1;
QY 7 ATAAEDFDVDPWVQVSSAKGIDYDKIYFSGSKIDKELINIRATGQRPHHPLRRGI 66
DB 38 ATAPTEVDVDPWVSSAKGIDYDKIYFSGSKIDKELINIRATGQRPHHPLRRGI 97
QY 67 FFSHRDMNOVLDAENKKPFYLYTGRGSPSEAMVGHLLPIFTKMLQDFNVPLVYIOMT 126

DB 98 FFSHRDLHLITLREBGRKFYLYTGRGSPSSGLHGHVLPFIMTKLOLQTFVPLVYIOLT 157
QY 127 DDEKYLKMDLTLDQAVGDAVENAKDIACGPDINKFIFSDLDYMGMSGGFYKNNVYIOK 186
DB 158 DDEKYLKMDLTLDQAVGDAVENAKDIACGPDINKFIFSDLDYMGMSGGFYKNNVYIOK 217
QY 187 HTFENOVKGIFFGTSDSDICISFPALQOAPSPFSNPFQIFRDRDIOCLIPALDODPY 246
DB 218 CVTFENOVKGIFFGTSDSDICISFPALQOAPSPFSNPFQIFRDRDIOCLIPALDODPY 276
QY 247 FFRTRDVAAPRIGYIPKALHSTFFPALQOAKTKMSADPNSSIFLDTAOKITKNNKA 306
DB 277 FFRTRDVAAPRIGYIPKALHSTFFPALQOAKTKMSADPNSSIFLDTAOKITKNNKA 336
QY 307 FSGGRDTEHROFGNCVDVSMYLTFFLEDDDKLEOIRKDYTSGAMLTGELKALIE 366
DB 337 FSGGRVTVVEHRLKGGVPEVDVSYOLKFFLEDDDALEVRVAVSKGEMLTGELKALIE 396
QY 367 VLOPLIAEHQARKEVTDEIVKEFMPRKLSP 398
DB 397 TLTPIVEHQARKLITDEVLDKYEFLRPLKF 428
RESULT 7
Q9VHG2 PRELIMINARY; PRT; 430 AA.
ID Q9VHG2
AC Q9VHG2
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Aats-trp protein.
GN AATS-TRP OR CG9735.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballieu K.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Bense P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M.C., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Galburt W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshireli A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclebo J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissenbach J.,
 Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 "The genome sequence of *Drosophila melanogaster*."
 Science 287:2185-2195(2000).
 EMBL: AEO03682; AAG22136.1; -
 DR EMBL: AB0010803; Aats-1p.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_1b.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PRO1039; TRNASYNTHTRP.
 DR TIGRFAMs: TIGR00233; trps; 1.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
 DR SEQUENCE 430 AA; 47971 MW; 2938EECC9E979F CRC64;

Query Match 64.8%; Score 1370.5; DB 5; Length 430;
 Best Local Similarity 64.5%; Pred. No. 5, 4e-109;
 Matches 253; Conservative 62; Mismatches 76; Indels 1; Gaps 1;

7 ATAEEDFVDPMTVOTSSAKGIDYDKLIVFGSSKIDKELINRERATGQRPHPFLRGI 66
 38 ATAPTEDVDVDPWNVASNDAGVDYDKLIRFGSSKIDDEILARFEKITGKPAHIFTRGM 97
 67 FFSHDMNOVLDAVENKKPFYLYTGRSPSEAMHGHILPFTKMLQOVENVPLVLT 126
 98 FFSHRDLHTLTLREGKPFYLYTGRSPSSGLHVLVPMTKWLQSTFVPLVLT 157
 127 DDEKYLMDLTLDOAGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKVVYK 186
 158 DDEKTLMDLTLDOAGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKVVYK 217
 187 HTFNQVKGIFGFTSDDCIGKISFPALQAPSFNSFPOIFRDRDIOCLIPCAIDQ 246
 218 CTFPNQVKGIFGFTSDDCIGKISFPALQAPSFNSFPOIFRDRDIOCLIPCAIDQ 276
 247 FPMTDVAPRIGVPPALHSTFPPALQAGOTKMSASDPSNLSFLLTPAKOIKTYNKA 306
 277 FPMTDVAPRIGVPPALHSTFPPALQAGOTKMSASDPSNLSFLLTPAKOIKTYNKA 336
 307 FSGGDTLIEHRQFGNCVDVDSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKALIE 366
 337 FSGGDTLIEHRQFGNCVDVDSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKALIE 396
 367 VLQPLIAHQARKKEVTDEIVKEFMTPRKLSF 398
 397 TLTPIVEHQARKKLITDEVLDKRYFELRPLKF 428

RESULT 8
 Q9SR15
 ID Q9SR15 PRELIMINARY; PRT; 402 AA.
 AC Q9SR15;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative tryptophanyl-tRNA synthetase.
 DE F7018.7 OR A13G04600.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
 RA Rensing C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC F7018 genomic sequence."
 RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

[2]
 RN SEQUENCE FROM N.A.
 RP Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.W., Quach H.L., Toriumi M., Yu G., Bowser L.,
 RA Carlini P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 RA Seki M., Shimizu P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.,
 RT "Arabidopsis Full Length cDNA Clones."
 RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC011437; AAF04890.1; -
 DR EMBL: AY080709; AAL85027.1; -
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_1b.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PRO1039; TRNASYNTHTRP.
 DR TIGRFAMs: TIGR00233; trps; 1.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
 DR Aminoacyl-tRNA synthetase.
 DR SEQUENCE 402 AA; 45754 MW; B9BEAV5E85D6CD15 CRC64;

Query Match 61.5%; Score 1301; DB 10; Length 402;
 Best Local Similarity 60.3%; Pred. No. 4, 4e-103;
 Matches 240; Conservative 66; Mismatches 88; Indels 4; Gaps 2;

6 DATEAE-EDFVDPMTVOTSSAKGIDYDKLIVFGSSKIDKELINRERATGQRPHPFLR 63
 7 DEREAESSEQVNVPEVSAKDGKIDYDKLIDFCQGLDLSLIDRVQRLTSRQHFELR 66
 64 RGIFFSHDMNOVLDAVENKKPFYLYTGRSPSEAMHGHILPFTKMLQOVENVPLV 123
 67 RSVFPAHRDPEHLLDAVERGDKPFYLYTGRSPSEAMHGHILPFTKMLQOVENVPLV 126
 124 QMTDDEKYLMDLTLDOAGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKVVYK 183
 127 QMTDDEKYLMDLTLDOAGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKVVYK 184
 184 IOXAVTNQVKGIFGFTSDDCIGKISFPALQAPSFNSFPOIFRDRDIOCLIPCAIDQ 243
 185 VQKCVTLNKGKIFGFTSDDCIGKISFPALQAPSFNSFPOIFRDRDIOCLIPCAIDQ 244
 244 DPFMTDVAAPRIGVPPALHSTFPPALQAGOTKMSASDPSNLSFLLTPAKOIKTYNKA 303
 245 DPFMTDVAAPRIGVPPALHSTFPPALQAGOTKMSASDPSNLSFLLTPAKOIKTYNKA 304
 304 KHAFFGDTLIEHRQFGNCVDVDSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKALIE 363
 305 KHAFFGDTLIEHRQFGNCVDVDSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKALIE 364
 364 LIEVLQPLIAHQARKKEVTDEIVKEFMTPRKLSF 401
 365 LIEVLQPLIAHQARKKEVTDEIVKEFMTPRKLSF 402

RESULT 9
 Q9UR12
 ID Q9UR12 PRELIMINARY; PRT; 324 AA.
 AC Q9UR12;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Y80D3A.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 NCBI_TaxID=6239;
 NCBI_TaxID=6239;
 RP SEQUENCE FROM N.A.
 RA Harris B.R.;
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99065613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
  investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: AL132853; CAB60439.1; -.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR002306; tRNA-synt_1b.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TENASYNTHTRP.
SQ SEQUENCE 324 AA; 36289 MW; 6E687E6E420ECC12 CRC64;

Query Match 48.5%; Score 1027; DB 5; Length 324;
Best Local Similarity 64.9%; Pred. No. 9.6e-80;
Matches 194; Conservative 39; Mismatches 64; Indels 2; Gaps 1;

QY 4 GPDATAEEDFVDPWTVQTSSAKGIDYDKLIVFGSSKIDKELINRIRATGQRPHHFLRG 63
DB 23 GGGVQDEDEBRVTPWETTTKATGIDYDKLIVFGCKDEDEIAREERTGKASPMR 82
QY 64 RGIFSHRDNNQVLDAVENKKPPLYTGRCPSSEAMVGHLPFTKMLQDVFNPLVIT 123
DB 83 RGMFFAHRDLTALIDRKEQGRPELYTGRGASSGSLHGLVPEFIETKMLQEVFVPLVI 142
QY 124 QMTDDEKYLKMDLTLDQAVGDVAENAKDIIACGPDINKTFISDLDYMGSSGFYKNVYK 183
DB 143 QMTDDEKYLKMDLVKAVGAKKMARENMKDIIISVGDPDKTFIFNNFD--MCPPEYEVYK 200
QY 184 IQKHTVFNQYKGFETDSDICKISFPALQAAAPSFNSPQIFRRTDIQCLIPCAIDQ 243
DB 201 IKKVVNTQNAARLFGFTPEDCGKAAPFAVEAAPCFASSPQIFGKRNIPCLIPCAIDQ 260
QY 244 DPFYFNRDVAPRIGYKPKPALHSTFPALQAGQTKMSADPNSSIFLDTAQIKTKV 302
DB 261 DPFYFNRDVAPRIGYKPKPALHSTFPALQAGQTKMSADPNSSIFLDTAQIKTKV 319

RESULT 10
08SQY5 PRELIMINARY; PRT; 385 AA.
AC 08SQY5;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Trypophanyl tRNA synthetase.
GN EC011 0530.
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryoniidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GB-M1;
RC Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=GB-M1;
MEDLINE=21576510; PubMed=11719906;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
  RA Premier G., Barbe V., Peyretailade E., Broctier P., Wincker P.,
  RA Delbec F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
  RA Weisenbach J., Vives C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
  Encephalitozoon cuniculi."
RL Nature 414:450-453(2001).
DR EMBL: AL590450; CAD25963.1; -.
SQ SEQUENCE 385 AA; 44188 MW; 28653AC95FCC859 CRC64;

Query Match 44.5%; Score 942; DB 5; Length 385;
Best Local Similarity 46.1%; Pred. No. 2.3e-72;
Matches 179; Conservative 76; Mismatches 123; Indels 10; Gaps 3;

```

```

QY 12 EDFVDPMVTQTSASAK-----GIDYDKLIVFGSSKIDKELINRIRATGQRPHHFLRG 67
DB 3 EQRTIPMDVEVYSTDEVPALIDYDKLINFGCEKRNQALADLEKLSGPAHYFFRRGIV 62
QY 68 FSHRDNNQVLDAVENKKPPLYTGRCPSSEAMVGHLPFTKMLQDVFNPLVITQMTD 127
DB 63 FAHRDNLNLLDEIANNRPPLYTGRGSSKTHIGHITIPFLCKYQDAFKRLVLIQTDE 122
QY 128 DEKYLKMDLTLDQAVGDVAENAKDIIACGPDINKTFISDLDYMGSSGFYKNVYK 187
DB 123 DEKFLKMSRLDQAVGRENKIDIVALGFDEKLYTIFSNV--ASHFEEENILKISKT 179
QY 188 VTFNQYKGFETDSDICKISFPALQAAAPSFNSPQIFRRTDIQCLIPCAIDDPYF 247
DB 180 ININEAIKFGFDMSSNICQGVFPAKEIPCCSSSRFTGK--GAMCLVPAVADDPFF 236
QY 248 RMTRDVAPRIGYKPKPALHSTFPALQAGQTKMSADPNSSIFLDTAQIKTKVKNKAF 307
DB 237 RLARDKAKALGKKPSIYVSLPDLKGVNRKMSADPNSSIFLDTAQIKTKVKNKAF 296
QY 308 SGGRTDIEHRQPGNCVDVSPFMTLTFLEDDDLKQIRKYTSGAMLGELKALIEV 367
DB 297 SGGRTDIEHRQPGNCVDVSPFMTLTFLEDDDLKQIRKYTSGAMLGELKALIEV 356
QY 368 LQPLIAHQARREVTDEIVKEFMTPRK 395
DB 357 IQEFSVRIOESKRKVTDDLRAFIDINK 384

RESULT 11
08U453 PRELIMINARY; PRT; 385 AA.
AC 08U453;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Trypophanyl tRNA synthetase.
GN PF0241.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.W., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE010149; AL80365.1; -.
KW Aminoacyl-tRNA synthetase; Complete proteome.
SQ SEQUENCE 385 AA; 45178 MW; 3A7A628958200CCC CRC64;

Query Match 39.1%; Score 826.5; DB 17; Length 385;
Best Local Similarity 45.9%; Pred. No. 1.8e-62;
Matches 180; Conservative 69; Mismatches 116; Indels 27; Gaps 10;

QY 11 EEDF-VDPWTVQTSSAKGIDYDKLIVFGSSKIDKELINRIRATGQRPHHFLRGIFPS 69
DB 2 EEEFKTPEVEGEG---IDYKLLIBQFSTPLDLDLERTALRLTSELPFPRKRFPS 57
QY 70 HRDNNQVLDAVENKKPPLYTGRCPSSEAMVGHLPFTKMLQDVFNPLVITQMTD 129
DB 58 HRDYNKVLDDYEQGKGFYLTGRGSG-PMHGIIIPFATKMLDEKFPVNLVYIQTDE 116
QY 130 KYLWMD-LTLDQAVGDVAENAKDIIACGPDINKTFISDLDYMGSSGFYKNVYK 188
DB 117 KFLFKNELTFEDTKYAWYQNIIDIIAVGDPDKTFIFQSSF---TKIYEMAIPIAKKI 172
QY 189 TFNQYKGFETDSDICKISFPALQAAAPSFNSPQIFRRTDIQCLIPCAIDDPYR 248
DB 173 NFSMAVAVGFTQSKIGMIFFPALQAAAPF-----FEKK---RCLIPAIADDPYR 222

```

249 MTRDVAPRIGYKPKPALHSTFPALOGAOTKMSASDPNSSIFLDTAKOIKTKVNHAS 308
 223 LORDAESIGYKTKALHSHKFPPLTGLEKMSASAFERAILTLDNPEAGKINKKAL 282
 309 GGRDTIEHROFGNCDDVVSFMYLTFPLEDDDKLEQIRKDY---TSGAMLTGELKALI 365
 283 GGQPTLKEORKEKGNPEKCVFPMLEIFEPDDK--KLMEYVACWNGELGCECKRYLI 340
 366 EYLQPLIAEHQARKEVTDIYKEFMTPKLS 397
 341 ORVOEFLKEHOKRRK-AEKLVEKFKYTGKLA 371

RESULT 12

Q9UIF5 PRELIMINARY; PRT; 490 AA.
 AC Q9UIF5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Tryptophanyl-tRNA synthetase.
 GN L1063.04.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_Taxid=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA Bothe G., Pohl T., Ivens A.C., Lawson D., Quail M., Rajandream M.A.,
 RA Barrell B.G.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA MEDLIN=98146435; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RL "A physical map of the Leishmania major Friedlin genome."
 RL Genome Res. 8:135-145(1998).
 DR EMBL; AL121862; CAB58393.1;
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002306; Ttp tRNA-synt_1b.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PRO1039; TRANSSYNTHTRP.
 DR PROSITE; PS00178; AA_tRNA_LIGASE_1; 1.
 KM Aminoacyl-tRNA synthetase.
 SQ SEQUENCE 490 AA; 54904 MW; 172C68622C511D3D CRC64;

Query Match 37.6%; Score 795.5; DB 5; Length 490;
 Best Local Similarity 38.1%; Pred. No. 1.2e-59;
 Matches 173; Conservative 87; Mismatches 129; Indels 65; Gaps 11;

QY 3 HGPDATEAEEDFVDPWTQTSAGKIDYDKLIVFGSSKID-----KELINRIERAT- 54
 DB 40 HPRDGGAGAEVITPVAVAKGPGQINRYVLTIFAERMDGARGHMDVAKCKRIM 99
 QY 55 -----GGRP-----HFLRGGIFSSHDM 73
 DB 100 TTTPSANAGVATPVLDAAVAGDALQAASAPMQRQOPPALALHFFHRDIAFSHRDL 159
 QY 74 NGVLDAVE-----NKKPFYLTGGRGSSSEAMVGHILPIFTKMLQDVNVPLVIOMTD 128
 DB 160 HRAALVDIESVKTGESVFLYTGGRGSGAGTMHGHVLPMLTKYLDVFSLPVLIQITD 219
 QY 129 ERYLWKDLTLIDQAVGDAV--ENAKDIIACGFDINKTFISDLDYMGSSGFYKAVVXKIOK 186
 DB 220 EKFLPRDVFPEAGKADLRISNLIKQIIAENFNPRHFIIRNTHYMG---DMVPTVARLQR 276
 QY 187 HYTFNOVKGIFGTSDDCIGKISFPALQAPSFNSFPQIFR--DRTDIQCLIPCAIDDD 244
 DB 277 SMTGNVAKHTLITGTDSDNVGKLAFPATOAPCFSTAFRRVLNRGDR-PWRCLIPCAIDDD 335

245 PYFRMTRDVAPRIGYKPKPALHSTFPALOGAOTKMSAS-DPNSSIFLDTAKOIKTKVN 303
 336 PPFVITRAALAIRKQLPPALHHTKPLPALKLEHMHSSAAEKGVTITHTDKQVRKL- 394
 304 KHAESGGRDTIEHROFGNCDDVVSFMYLTFPLEDDDKLEQIRKDYTSGAMLTGELKKA 363
 395 BRAFGGCGATLEQOMQETGANIELDVAOYLRFPCDDTLFADVNTQRRYSGTLNLSGEVDL 454
 364 LIE-VLQPLIAEHQARKEVTDIYKEFMTPKL 396
 455 AADCIIRVLDHWRRERRATVTDVDDVEFCRIIDI 488

RESULT 13

Q95295 PRELIMINARY; PRT; 136 AA.
 AC Q95295;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (Fragment).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_Taxid=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SMALL INTESTINE;
 RA Winzore A.K., Fredholm M., Davies W.;
 RT "Evaluation and characterization of a porcine small intestine cDNA
 RT library."
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z81267; CAB03585.1;
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR002306; Ttp tRNA-synt_1b.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PRO1039; TRANSSYNTHTRP.
 FT NON_TER 1 1
 FT NON_TER 136 136
 SQ SEQUENCE 136 AA; 15404 MW; 951F75D1B9CD0617 CRC64;

Query Match 32.2%; Score 682; DB 6; Length 136;
 Best Local Similarity 94.1%; Pred. No. 1e-50;
 Matches 128; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 211 PAIQAPSFNSFPQIFDRDIDQCLIPCAIDDDPYFRMTRDVAPRIGYKPKALHSTFF 270
 DB 1 PAIQAPSFNSFPQIFDRDIDQCLIPCAIDDDPYFRMTRDVAPRIGYKPKALHSTFF 60
 QY 271 PALOGAOTKMSASDPNSSIFLDTAKOIKTKVNHASFGGRPTIEHROFGNCDDVVSF 330
 DB 61 PALOGAOTKMSASDPNSSIFLDTAKOIKTKVNHASFGGRPTIEHROFGNCDDVVSF 120
 QY 331 MYLTFPLEDDDKLEQI 346
 DB 121 MYLTFPLEDDDKLEQI 136

RESULT 14

O8TYF7 PRELIMINARY; PRT; 374 AA.
 AC O8TYF7;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Tryptophanyl-tRNA synthetase.
 GN TRPS OR MK0343.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyrii; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_Taxid=320;
 RN [1]
 RP SEQUENCE FROM N.A.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:26:29 ; Search time 8.74909 Seconds

(without alignments)
1900.998 Million cell updates/sec

Title: US-09-813-718-10_COPY_71_471

Perfect score: 2116
Sequence: 1 SMHGDATAEEDFVDPWTV.....VNDIVKERTMRKLSFDDQ 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2101	99.3	471	1	SYN_HUMAN
2	2028.5	95.9	475	1	SYN_BOVIN
3	1961.5	92.7	475	1	SYN_RABIT
4	1938	91.6	481	1	SYN_MOUSE
5	1210	57.2	395	1	SYN_SCHPO
6	1163	55.0	432	1	SYN_YEAST
7	910.5	43.0	381	1	SYN_SULTO
8	905	42.8	380	1	SYN_SULSO
9	803	37.9	385	1	SYN_PYRAB
10	796	37.6	386	1	SYN_PYRHO
11	735	34.7	375	1	SYN_PYRAE
12	534.5	25.3	380	1	SYN_HALNI
13	409.5	19.4	370	1	SYN_MERVA
14	397.5	18.8	364	1	SYN_MERTH
15	386	18.2	134	1	SYN_ENCCU
16	370.5	17.5	420	1	SYN_ARCFU
17	353	16.7	374	1	SYN_AERPE
18	299.5	14.2	426	1	SYN_THERVO
19	268.5	12.7	426	1	SYN_THERAC
20	192	9.1	323	1	SYN_ARCFU
21	185.5	8.8	341	1	SYN_STRPN
22	182.5	8.6	340	1	SYN_STRPY
23	175.5	8.3	341	1	SYN_CLODO
24	174.5	8.2	341	1	SYN_LACIA
25	172	8.1	394	1	SYN_YEAST
26	170.5	8.1	366	1	SYN_SUSO
27	169	8.0	395	1	SYN_AQUIA
28	166.5	7.9	346	1	SYN_AQUAR
29	163.5	7.7	343	1	SYN_CHLTR
30	159.5	7.5	337	1	SYN_BORBU
31	158	7.2	346	1	SYN_TREPA
32	152	7.2	344	1	SYN_CHLPM
33	146	6.9	326	1	SYN_HELPY

34	138	6.5	319	1	SYN_MERTH	027795 methanobact
35	136.5	6.5	528	1	SYN_BOVIN	029465 bos taurus
36	135.5	6.4	328	1	SYN_THERMA	094722 thermocoga
37	133	6.3	328	1	SYN_BACST	P00953 bacillus st
38	132.5	6.3	343	1	SYN_MYCLE	Q49901 mycobacteri
39	132.5	6.3	528	1	SYN_HUMAN	P54577 homo sapien
40	130	6.1	326	1	SYN_HELPY	092124 heliocobact
41	129	6.1	401	1	SYN_SCHPO	014055 schizosacch
42	128.5	6.1	379	1	SYN_YEAST	P04803 saccharomyc
43	128	6.0	334	1	SYN_HAEIN	P43835 haemophilus
44	124.5	5.9	347	1	SYN_MYCGE	P47372 mycoplasma
45	122	5.8	350	1	SYN_CAEEEL	P46579 caenorhabdi

ALIGNMENTS

RESULT 1	SYN_HUMAN	STANDARD;	PRT;	471 AA.
AC	P23381; P78535; Q9UDJ3;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)			
DE	(TFPRS) (TFP53) (HMR5).			
GN	WARS.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92105071; PubMed=1761529;			
RA	Rubin B.Y., Anderson S.L., Xing L., Powell R.J., Tate W.P.;			
RT	"interferon induces tryptophanyl-tRNA synthetase expression in human			
RT	fibroblasts.";			
RL	J. Biol. Chem. 266:24245-24248(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92107982; PubMed=1763065;			
RA	Flechner J., Rasmussen H.H., Justesen J.;			
RT	"Human interferon gamma potentially induces the synthesis of a 55-kDa			
RT	protein (gamma 2) highly homologous to rabbit peptide chain release			
RT	factor and bovine tryptophanyl-tRNA synthetase.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:11520-11524(1991).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92112058; PubMed=1765274;			
RA	Frolova L.Y., Sudomoina M.A., Grigorjeva A.Y., Zinovjeva O.L.,			
RA	Kiselev L.V.;			
RT	"Cloning and nucleotide sequence of the structural gene encoding for			
RT	human tryptophanyl-tRNA synthetase.";			
RL	Gene 109:291-296(1991).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92164636; PubMed=1537332;			
RA	Buwallt U., Flohr T., Boettger B.C.;			
RT	"Molecular cloning and characterization of an interferon induced			
RT	human cDNA with sequence homology to a mammalian peptide chain			
RT	release factor.";			
RL	EMBO J. 11:489-496(1992).			
RN	[5]			
RP	SEQUENCE OF 1-13 FROM N.A.			
RX	MEDLINE=96319994; PubMed=8724762;			
RA	Sokolova I.V., Narovianskii A.N., Amchenkova A.M., Turpaev K.T.;			
RT	"Alternative splicing of 5'-terminal exons of the human tryptophanyl-			
RT	tRNA synthetase gene.";			
RL	Mol. Biol. (Mosk) 30:319-329(1996).			
RN	[6]			
RP	SEQUENCE OF 1-141 AND 182-471 FROM N.A.			
RC	TISSUE=Spem;			
RX	MEDLINE=93292992; PubMed=7685728;			

RA Prolova L.Y., Grigorjeva A.Y., Sudomoina M.A., Kisselev L.L.;
 RT "The human gene encoding tryptophanyl-tRNA synthetase: interferon-
 response elements and exon-intron organization.";
 RL Gene 128:237-245(1993).
 RN [7]
 RP SEQUENCE OF 265-276; 278-296; 299-317 AND 350-365.
 RC TISSUE=Keratinocytes;
 RX MEDLINE=93162043; Pubmed=1286667;
 RA Rasmussen H.H., van Damme J., Puyse M., Geeser B., Celis J.E.,
 "Van dekerckhove J.";
 RT "Interferon-induced protein with release factor activity is a
 tryptophanyl-tRNA synthetase.";
 RL Protein database of normal human epidermal keratinocytes.";
 RT Electrophoresis 13:960-969(1992).
 RN [8]
 RP FUNCTION
 RX MEDLINE=9225128; Pubmed=1373391;
 RA Bange F.-C., Flohr T., Buwitt U., Boettger E.C.;
 RT "An interferon-induced protein with release factor activity is a
 tryptophanyl-tRNA synthetase.";
 RL FEBS Lett. 300:162-166(1992).
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- INDUCTION: BY INTERFERON GAMMA.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 or send an email to license@ebi.ac.uk).

SQ SEQUENCE 471 AA; 53165 MW; E96344449053A0D0 CRC64;
 Query Match 99.3%; Score 2101; DB 1; Length 471;
 Best Local Similarity 99.5%; Pred. No. 2,8e-165;
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SNHGPGATEAEDSFVDPMTVQTSASAKIDYDKLIYVFGSSKIDKELINERATGQRPHH 60
 DB 71 SNHGPGATEAEDSFVDPMTVQTSASAKIDYDKLIYVFGSSKIDKELINERATGQRPHH 130
 QY 61 FLNRGIFFSHRDMNQVLADAYENKPPVLYTGRSPSSAMVGHILPIFTKMLQDVFNVP 120
 DB 131 FLNRGIFFSHRDMNQVLADAYENKPPVLYTGRSPSSAMVGHILPIFTKMLQDVFNVP 190
 QY 121 LVIQMTDDERKYLKMDLTLDQAAGDAVENAKDIIACGFDINKTPIFSDLDYMGSSGFYKN 180
 DB 191 LVIQMTDDERKYLKMDLTLDQAAGDAVENAKDIIACGFDINKTPIFSDLDYMGSSGFYKN 250
 QY 181 VKKIQHVFENQVKGIFGFTSDSDICKISFPALQAPFSNSFPQIFRDRDIOCLIPCA 240
 DB 251 VKKIQHVFENQVKGIFGFTSDSDICKISFPALQAPFSNSFPQIFRDRDIOCLIPCA 310
 QY 241 IDDPYFRMTRDVAPRIGYKPKALHSTFPFALQAGTQMSADPNSSIFLDTAKQIKT 300
 DB 311 IDDPYFRMTRDVAPRIGYKPKALHSTFPFALQAGTQMSADPNSSIFLDTAKQIKT 370
 QY 301 KVNKAHPSGGRDTEHRQFGNCQDVSPMLTFLLEDKLEQIKRKYTSGAMLTGEL 360
 DB 371 KVNKAHPSGGRDTEHRQFGNCQDVSPMLTFLLEDKLEQIKRKYTSGAMLTGEL 430
 QY 361 KKALIEVLOPLIAEHQARKEVTDIVKEPMTPRKLSFPDQ 401
 DB 431 KKALIEVLOPLIAEHQARKEVTDIVKEPMTPRKLSFPDQ 471

RESULT 2
 SYM_BOVIN
 ID SYM_BOVIN STANDARD; PRT; 475 AA.
 AC P17248;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
 DE (TrpRS).
 GN WARS.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Retina;
 RX MEDLINE=91329348; Pubmed=1907847;
 RA Garret M., Pajot B., Trezeguet V., Labouesse J., Merle M.,
 Gandar J.-C., Benedetto J.-P., Sallatranque M.-L., Alterio J.,
 Gueguen M., Sarger C., Labouesse B., Bonnet J.;
 RA "A mammalian tryptophanyl-tRNA synthetase shows little homology to
 RT prokaryotic synthetases but near identity with mammalian peptide
 RT chain release factor.";
 RL Biochemistry 30:7809-7817(1991).
 RN [2]
 RP SEQUENCE OF 17-475 FROM N.A.
 RC TISSUE=Pancreas;
 RA Garret M., Trezeguet V., Pajot B., Gandar J.-C., Merle M.,
 Gueguen M., Benedetto J.-P., Sarger C., Alterio J., La Bouesse B.,
 Labouesse J., Bonnet J.;
 RA Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.
 RL -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X53918; CAA37872.1; -
 DR EMBL: X52113; CAA36356.1; -
 DR PIR: A40279; YMBQ.
 DR PIR: S14540; S14540.
 DR InterPro: IPR002306; Trp-trna-synt_1b.
 DR InterPro: IPR0002306; WHEP-TRS.
 DR InterPro: IPR002305; trna-synt_1b.
 DR InterPro: IPR001412; trna-synt_1.
 DR Pfam: PF00458; WHEP-TRS; 1.
 DR Pfam: PF00579; trna-synt_1b; 1.
 DR PRINTS: PRO1039; TRNASYNTHTRP.
 DR TIGRPFAM: TIGR00233; trps; 1.
 DR PROSITE: PS00178; AA-TRNA-LIGASE_1; 1.
 DR PROSITE: PS00762; WHEP-TRS; 1.
 DR KMW: Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT DOMAIN 24 69 WHEP-TRS.
 FT SITE 169 177 DISPENSABLE TO THE CATALYTIC ACTIVITY.
 FT SITE 353 357 "KMSKS" REGION.
 FT CONFLICT 17 17 L -> M (IN REF. 2).
 SQ SEQUENCE 475 AA; 53729 MW; 37E531750137EB32 CRC64;

Query Match 95.9%; Score 2028.5; DB 1; Length 475;
 Best Local Similarity 95.8%; Pred. No. 2.6e-159;
 Matches 384; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 1 SNHGPDATAEEDFDVDPMTVOTSSAKGIDYDKLIYRFGSSKIDKELNIRIERATGQRPH 60
 DB 76 SGEGLDADAEDEDFVDPMVTQVSSAKGIDYDKLIYRFGSSKIDKELNIRIERATGQRPH 135
 QY 61 FLRRGIFSHRDNMQVLDAYENKRPYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 120
 DB 136 FLRRGIFSHRDNMQVLDAYENKRPYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 195
 QY 121 LVIQMTDDEKYLKQLTLDQAYGDAVENAKDIIACGPDINTKFIETSDLDYMGSSGFYKN 180
 DB 196 LVIQMTDDEKYLKQLTLDQAYGDAVENAKD-IITGSPDINTKFIETSDLDYMGSSGFYKN 254
 QY 181 VVKIQHTVFNQVKGIFGFTSDCIKISFPALQAPFSNSFPQIFRDRDIIQCLIPCA 240
 DB 255 VVKIQHTVFNQVKGIFGFTSDCIKISFPALQAPFSNSFPQIFRDRDIIQCLIPCA 314
 QY 241 IDDPYFRMTRDVAPRIGYPRKPALHSTFPPALQAGQTKMSASDPNSSIFLTDYAKQIKT 300
 DB 315 IDDPYFRMTRDVAPRIGYPRKPALHSTFPPALQAGQTKMSASDPNSSIFLTDYAKQIKT 374
 QY 301 KKNKAFFSGGRPTIEHROPGNCVDVSPMYLTFLEDDDLKQIRKRYTGAMLTGEL 360
 DB 375 KKNKAFFSGGRPTIEHROPGNCVDVSPMYLTFLEDDDLKQIRKRYTGAMLTGEL 434
 QY 361 KKALEVLQPLIAEHQARRKEVTDIVKEFMTPRKLSFDFQ 401
 DB 435 KKALEVLQPLIAEHQARRKEVTDIVKEFMTPRKLSFDFQ 475

RESULT 3
 ID SYM RABIT STANDARD; PRT; 475 AA.
 AC P23612; Q28607;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 15-NOV-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TrpRS).

GN WARS.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90239043; PubMed=2185472;
 RA Lee C.C., Craigen W.J., Murny D.M., Harlow E., Caeky C.T.;
 RT "Cloning and expression of a mammalian peptide chain release factor
 RT with sequence similarity to tryptophanyl-tRNA synthetases.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3508-3512(1990).
 RN [2]
 RP REVISIONS TO 169-174 AND 227-228, AND FUNCTION.
 RX MEDLINE=9409008; PubMed=840467.
 RA Frolova L.Y., Dalphin M.E., Justesen J., Powell R.J., Drugeon G.,
 RA McCaughan K.K., Kisselev L.L., Tate W.P., Haenni A.-L.;
 RT "Mammalian polypeptide chain release factor and tryptophanyl-tRNA
 RT synthetase are distinct proteins.";
 RL EMBO J. 12:4013-4019(1993).
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC dihydrophate + L-tryptophanyl-tRNA(Trp).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.
 CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE AN EUKARYOTIC
 CC RELEASE FACTOR (ERF).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M33460; AAA31246.1; ALT_SEQ.
 DR EMBL: U02595; AAB60257.1; -
 DR PIR: A35904; YMRBPR.
 DR InterPro: IPR002306; Trp-trna-synt_1b.
 DR InterPro: IPR000738; WHEP-TRS.
 DR InterPro: IPR002305; trna-synt_1b.
 DR InterPro: IPR001412; trna-synt_1.
 DR Pfam: PF00458; WHEP-TRS; 1.
 DR Pfam: PF00579; trna-synt_1b; 1.
 DR PRINTS: PRO1039; TRNASYNTHTRP.
 DR TIGRPFAM: TIGR00233; trps; 1.
 DR PROSITE: PS00178; AA-TRNA-LIGASE_1; 1.
 DR PROSITE: PS00762; WHEP-TRS; 1.
 DR KMW: Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT DOMAIN 23 68 WHEP-TRS.
 FT SITE 168 177 "HIGH" REGION.
 FT SITE 353 357 "KMSKS" REGION.
 SQ SEQUENCE 475 AA; 53729 MW; 33BC9E718FFA5DC4 CRC64;

Query Match 92.7%; Score 1961.5; DB 1; Length 475;
 Best Local Similarity 91.5%; Pred. No. 8.5e-154;
 Matches 367; Conservative 20; Mismatches 13; Indels 1; Gaps 1;

QY 2 NHG-DATEAEEDFDVDPMTVOTSSAKGIDYDKLIYRFGSSKIDKELNIRIERATGQRPH 60
 DB 75 SHGDEAVADKEDFDVDPMTVOTSSAKGIDYDKLIYRFGSSKIDKELNIRIERATGQRPH 134
 QY 61 FLRRGIFSHRDNMQVLDAYENKRPYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 120
 DB 135 FLRRGIFSHRDNMQVLDAYENKRPYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 194
 QY 121 LVIQMTDDEKYLKQLTLDQAYGDAVENAKDIIACGPDINTKFIETSDLDYMGSSGFYKN 180
 DB 195 LVQMSDEKYLKQLTLDQAYGDAVENAKDIIACGPDINTKFIETSDLDYMGSSGFYKN 254
 QY 181 VVKIQHTVFNQVKGIFGFTSDCIKISFPALQAPFSNSFPQIFRDRDIIQCLIPCA 240

```

Db 255 VKI0KVTN0YKGIFFGFTDSDICIGISFPALQAAFSFNSPQIFHGADIQCLIPCA 314
QY 241 IDDDPYFRMTRDVAPRIGYKPKPALHSTFPALQAGATKMSADPNSSIFLTDTAKQIKT 300
Db 315 IDDDPYFRMTRDVAPRIGYKPKPALHSTFPALQAGATKMSADPNSSIFLTDTAKQIKT 374
QY 301 KVNKHAFFSGGRDTEEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGEL 360
Db 375 KVNKHAFFSGGRDTEEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGEL 434
QY 361 KKALEVLQPLIAEHQARRKVTDEIVKEFMTPRKLSFDFQ 401
Db 435 KKEILIDVLQPLVAEHQARRKVTDEIVKEFMTPRQLCFHQ 475

RESULT 4
SYM_MOUSE STANDARD; PRT; 481 AA.
ID SYM_MOUSE
AC P32521;
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
DE WARS OR WRS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=95018226; PubMed=7932716;
RA Pajot B., Sarger C., Bonnet J., Garret M.;
RT "An alternative splicing modifies the C-terminal end of tryptophanyl-
RT tRNA synthetase in murine embryonic stem cells.";
RL J. Mol. Biol. 242:599-603(1994).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA (Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC short form; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: The short isoform is widely expressed, the
CC long form is found only in embryonic stem cells.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X69656; CAA49347.1; -
CC EMBL; X69657; CAA49348.1; -
CC PIR; S31461; S31461.
CC PIR; S31462; S31462.
CC MGD; MGI:104630; WARS.
CC InterPro; IPR002306; Trp tRNA-synt_1b.
CC InterPro; IPR000738; WHEP-TRS.
CC InterPro; IPR002305; tRNA-synt_1b.
CC InterPro; IPR001412; tRNA-synt_1.
CC Pfam; PF00458; WHEP-TRS; 1.
CC Pfam; PF00579; tRNA-synt_1b; 1.
CC PRINTS; PRO1039; TRNASYNTHTRP.
CC TIGRPFAMS; TIGR00233; trps; 1.
CC PROSITE; PS00178; AA tRNA LIGASE_1; 1.
CC PROSITE; PS00762; WHEP-TRS; 1.
CC Anticodon-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.

```

```

FT SITE 353 357 "KMSKS" REGION.
FT VARSLIC 476 481 MISSING (IN ISOFORM 2).
SQ SEQUENCE 481 AA; 54282 MW; B05M452C08074F52 CR664;
Query Match 91.6%; Score 1938; DB 1; Length 481;
Best Local Similarity 90.8%; Pred. No. 7,4e-152;
Matches 363; Conservative 21; Mismatches 16; Indels 0; Gaps 0;
QY 2 NHCPLATAEADFPVDVPTVQTSKAKIDYDKLIVRGSSKIDKELINRIERATGQRPPIHF 61
Db 76 NCDSDATKASEDFVDPWTVRTSSAKIDYDKLIVQGSSEKIDKELINRIERATGQRPPIHF 135
QY 62 LRGGIFFSRDMNQVLDAVENKKPFYLYTGRGSPSSAMVGHILPFIPTKMLQDVENVVL 121
Db 136 LRGGIFFSRDMNQVLDAVENKKPFYLYTGRGSPSSAMVGHILVPIPTKMLQDVENVVL 195
QY 122 VIQMTDEKRYLWKDLTLQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFEYKAV 181
Db 196 VIQMSDEKRYLWKDLTLQAYGYVENAKDIIACGFDINKTFIFSDLEVMGQSPGEYKAV 255
QY 182 VKI0KVTN0YKGIFFGFTDSDICIGISFPALQAAFSFNSPQIFRDRTDIQCLIPCAI 241
Db 256 VKI0KVTN0YKGIFFGFTDSDICIGISFPALQAAFSFNSPQIFRDRTDIQCLIPCAI 315
QY 242 DDDPYFRMTRDVAPRIGYKPKPALHSTFPALQAGATKMSADPNSSIFLTDTAKQIKT 301
Db 316 DDDPYFRMTRDVAPRIGYKPKPALHSTFPALQAGATKMSADPNSSIFLTDTAKQIKT 375
QY 302 VKNKHAFFSGGRDTEEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELK 361
Db 376 VKNKHAFFSGGRDTEEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELK 435
QY 362 KKALEVLQPLIAEHQARRKVTDEIVKEFMTPRKLSFDFQ 401
Db 436 KKEILIDVLQPLVAEHQARRKVTDEIVKEFMTPRQLCFHQ 475

RESULT 5
SYM_SCHRO STANDARD; PRT; 395 AA.
ID SYM_SCHRO
AC Q09692;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA
DE ligase) (TrpRS).
DE GN SPAC2F7.13C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP MEDLINE=1859360;
RP STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor A., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones I., Jones M., Leather S., McDonald J., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Ocell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Wolcraet G., Aert R., Robben J., Gymnopoulos B.,
RA Welter J., Vanstreels E., Rieger M., Schaefer M., Meller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

```

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Jelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revelante J.L., Moreno S., Armstrong J., Forburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakowski G.V., Useary D., Barrell B.G., Nurse P.,
 RT "The genome sequence of *Schizosaccharomyces pombe*,"
 RL Nature 415:871-880(2002).
 CC -1- CATABOLIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z50142; CAA90500.1; -
 DR InterPro; IPR002306; Trp_tRNA-synt_1b.
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PR01039; TRNASYNTHRP.
 DR TIGRFAMs; TIGR00233; trpS; 1.
 DR PROSITE; PS00178; AA tRNA_LIGASE_I; 1.
 KM Hypothetical protein; Aminoacyl-tRNA synthetase; Protein biosynthesis;
 KM Ligase; ATP-binding.
 FT SITE 91 100 "HIGH" REGION.
 FT SITE 275 279 "KMSKS" REGION.
 SQ SEQUENCE 395 AA; 44910 MW; E656AEBB76C5F9F9 CRC64;

Query Match 57.2%; Score 1210; DB 1; Length 395;
 Best Local Similarity 59.6%; Pred. No. 4.2e-92;
 Matches 235; Conservative 60; Mismatches 91; Indels 8; Gaps 4;

QY 11 EEDFDPVPTVQTS---SAKGIDYDKLYRFGSSKIDKELINIERATQGRPHHFRGCI 66
 DB 4 EEOIYTPMVVKGSIYDGEKGIYERLIYQFGTRKTPQLERFEKLTGKKPHLLRRGA 63
 QY 67 FFSHDMNOVLDAVENKKPFYLYTGGSPSSAMHVGHLIPFTKMLQVNPVPLVQMT 126
 DB 64 FFSHDFDMTILDRYEKKPFYLYTGGSPSSAMHVGHLIPFTKMLQVNPVPLVQMT 123
 QY 127 DDEKYLMD-LTLDDQAYGDAVENAKDIIACGPDINTFIIFSLDLYMGSSGFYKVVVKIQ 185
 DB 124 DDEKFLFKQGVSLDQGRFARENAKDIIAVGPDPKTFFIMNSTVYG--GAFYQVNVRIA 181
 QY 186 KAVTNQVKGIFGFTSDICIGKISPAIDAPSFNSFPQIRDRDIOCLIPCAIDDP 245
 DB 182 KCIITNQSACGFTSDISIGKIHFAISIQAPSFNSFPHINGAADIPCLIPCAIDDP 241
 QY 246 YRMRTRDVAIPRGYPKPAALHSFPFALGAGOTKMSASPNSSIFLTDFAKOIKTKVNG 305
 DB 242 YRLTRDVSGLKFKKPAALHSFPFALGQPSKMSASOSIAIFMTDIPNKKIKKNINHA 301
 QY 306 AFSGRDITIEHRQFQCDVVSFMYLTFPLEDDDKLQIRKDYTGAMLTGELKALI 365
 DB 302 AFSGGATIEIHRKQGNPDVVAQYLSFPLDDDKLKQLVNTYAAGTLSTGEMKGECI 361
 QY 366 EYLOPLIAEHQARKKAYDEIYKEFMT-PRKLSF 398
 DB 362 KLIQFVSDFOAARSKYDEATLDMFWDGSRKLEW 395

RESULT 6
 ID SYNC YEAST STANDARD; PRT; 432 AA.
 AC Q12109;
 DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Tryptophanyl-tRNA synthetase, cytoplasmic (EC 6.1.1.2) (tryptophan--
 DE tRNA ligase) (TrpRS).
 GN WRS1 OR YOL097C OR HRE432.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_Taxid=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96076631; PubMed=7502582;
 RA Vandenbol M., Durand P., Portetelle D., Hilger F.;
 RT "Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV
 RT including the Trp-H3 retrotransposon, the sufl1(+) frameshift
 RT suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-1a and a
 RL delta element.";
 RL Yeast 11:1069-1075(1995).
 RN [2]
 RP FUNCTION.
 RX MEDLINE=97197969; PubMed=9046085;
 RA John T.R., Ghosh M., Johnson J.D.;
 RT "Identification and expression of the *Saccharomyces cerevisiae*
 RT cytoplasmic tryptophanyl-tRNA synthetase gene.";
 RL Yeast 13:37-41(1997).
 CC -1- CATABOLIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z48149; CAA88164.1; -
 DR EMBL; Z74839; CAA99110.1; -
 DR SGD; S0005457; WRS1.
 DR InterPro; IPR002306; Trp_tRNA-synt_1b.
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PR01039; TRNASYNTHRP.
 DR TIGRFAMs; TIGR00233; trpS; 1.
 DR PROSITE; PS00178; AA tRNA_LIGASE_I; 1.
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT SITE 111 120 "HIGH" REGION.
 FT SITE 295 299 "KMSKS" REGION.
 SQ SEQUENCE 432 AA; 49350 MW; C408F169737E9736 CRC64;

Query Match 55.0%; Score 1163; DB 1; Length 432;
 Best Local Similarity 54.8%; Pred. No. 3.5e-88;
 Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

QY 7 ATEAEEDFVDPVTV-----QTSAGKIDYDKLYRFGSSKIDKELINIERATQGRPHHFR 61
 DB 19 STDVKEQVTVDPVGVGVGDEQGRAGNIDYDKLIFGTRPVNEETLKPFKQVYTGSRPHH 78
 QY 62 LRKGFSSHRDMNOVLDAVENKKPFYLYTGGSPSSAMHVGHLIPFTKMLQVNPVPL 121
 DB 79 LRKGLFSEBRDFTKLLDIEQKPFYLYTGGSPSSAMHVGHLIPFTKMLQVNPVPL 138
 QY 122 VIQMTDDEKYLMD-LTLDDQAYGDAVENAKDIIACGPDINTFIIFSLDLYMGSSGFYK 180
 DB 139 VIELTDEKFLFKKLTJINDVKNFARENAKDIIVGPDPKTFFISDIOYWG--GAFYET 196
 QY 181 VKIQKAVTNQVKGIFGFTSDICIGKISPAIDAPSFNSFPQIRDRDIOCLIPCA 240
 DB 197 VVRVSRQITGSAKAVFQFNDSDCIGKHFASIQATAFPSFPVVLGLPKTCLIPCA 256

```

QY 241 IDDDYFMRTRDVPAPRIGYKPAALHSTFPALQGAOTKMSADPNSIFLTDPAKQIK 300
DB 257 IDDDYFMRTRDVPADKLSKPLHSTRFPALQGAOTKMSADPNSIFLTDPAKQIK 316
QY 301 KVKKHAFFSGGRDPIEHRGFGNCDDVVSFMYLTFPLEDDDKLEQIRKDYTGAMLTGEL 360
DB 317 KINKAFVSGGQVADLHRELQGNPDVAVAYQISFFPDVDVFLKECDYKXSGELLSGEM 376
QY 361 KKAILEVLOPLIAEHQARKREVTDEIVKEFMTPRKL 396
DB 377 KKLCTETLOEFKAFQERRAQVDEETLTKFMPVPHKL 412

RESULT 7
SYM_SUITO STANDARD; PRT; 381 AA.
ID SYM_SUITO STANDARD; PRT; 381 AA.
AC O976M1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR ST0169.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=JCM 10545 / ?;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Akai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Oguchi A.,
RA Aoki K.-I., Maeda S., Yanagi M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain?";
RL DNA Res. 8:123-140(2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AEO00981; BAB65126.1; ALT INIT.
CC InterPro; IPR002306; Trp_cRNA-synt_1b.
CC InterPro; IPR002305; tRNA-synt_1b.
CC Pfam; PF00579; tRNA-synt_1b; 1.
CC TIGRPFAM; TIGR00233; trps_1.
CC PROSITE; PS00178; AA_tRNA_LIGASE_I; FALSE NEG.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC SITE 82 90 "HIGH" REGION.
CC SITE 254 258 "KMSKS" REGION.
CC SEQUENCE 381 AA; 44718 MW; 109F5A56AF7D7159 CRC64;

Query Match 43.0%; Score 910.5; DB 1; Length 381;
Best Local Similarity 50.5%; Pred. No. 1,7e-67;
Matches 169; Conservative 63; Mismatches 99; Indels 23; Gaps 9;
QY 12 EBD-VDPWYQTSAGK-IDYDKLYRFGSSKIDKELINIEKATGCRPHPLRGRIFFS 69
DB 3 ODENVTPWEV-----GKKVDYDKLIYQFGTKITSEIKKIKSIINDELHVMRLRDVFFS 57

```

```

QY 70 HRDMQVLDAYENKKPFYLYTGRSPSSAMHGHILPFIETKMLQDVENVPIQWOTDE 129
DB 58 HRDLVDLVQDDYDDGKGFYLYTGRAPSL-GMHLGLPFIETKMLQDVENVPIQWOTDE 116
QY 130 KYLMK-DLTLDQAYGAVENAKDIIACGFDINKTEIFSDLDYMGSSGFGYKVVKIQKHV 188
DB 117 KEMRBNREYTLDDTQWAVYNNIDIIAVGPNPKTFIFQOTEXI-----RMNYPIAIKAKKL 173
QY 189 TENQKGIQGFSDSCIGKISFPALQGAOTKMSADPNSIFLTDPAKQIKTRVNHAPS 248
DB 174 TFSBVARATGADTSSNIGIWMYPAQIAET-----MEKKE--RCLIPAGIDDPYWR 223
QY 249 MTRDVAFRIGYKPAALHSTFPALQGAOTKMSADPNSIFLTDPAKQIKTRVNHAPS 308
DB 224 LQDRIAESLGYKRAQIHKFLPPLTGPBGKSSSGPEFALYLTDDPKTVERKIMKVAS 283
QY 309 GGRDPIEHRGFGNCDDVVSFMYLTFPLE-DDDKLEQIRKDYTGAMLTGELKALIEV 367
DB 284 GGQPTIELRKRYGKNPDIVSFQWLYMFPEDPNKIKIEEDYRSGALLTGELKQITLEK 343
QY 368 LQPLIAEHQARKKE 381
DB 344 LNDPFEHRQKREE 357

RESULT 8
SYM_SUITO STANDARD; PRT; 380 AA.
ID SYM_SUITO STANDARD; PRT; 380 AA.
AC O972X0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR SSO0452.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awaysz M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtiss B.A.,
RA De Moers A., Erasmo G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Thieriault C., Tolstrup N.,
RA Charlebois R.L., Doocittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AEO06677; AAK40778.1; ALT INIT.
CC InterPro; IPR002306; Trp_cRNA-synt_1b.
CC InterPro; IPR002305; tRNA-synt_1b.
CC Pfam; PF00579; tRNA-synt_1b; 1.
CC PRINTS; PRO1039; TRNASYNTHTRP.
CC TIGRPFAM; TIGR00233; trps_1.
CC PROSITE; PS00178; AA_tRNA_LIGASE_I; FALSE NEG.

```

KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 FT SITE 81 89 "HIGH" REGION.
 FT SITE 253 257 "KMSKS" REGION.
 SQ SEQUENCE 380 AA; 44691 MW; CF8344CF63883680 CRC64;
 Query Match 42.8%; Score 905; DB 1; Length 380;
 Best Local Similarity 48.7%; Pred. No. 4.8e-67;
 Matches 190; Conservative 68; Mismatches 108; Indels 24; Gaps 10;

QY 11 EEDFVDPMTVQTSSAKG-IDYKLVRFSSSKIDKELINRIERATGQRPHPHPLRGIFFS 69
 DB 3 DEFTVTPMEVEG---KGKVDYDKLIVQFGQTKTEELKQRIKMLAGDL-HVMLRRVYFSS 56
 QY 70 HRDMQVLDAYENKKEFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVLVIQMTDDE 129
 DB 57 HRDLVDLVNDYKESKQFPLYTGAPSL-QGHIGHILPFTKMLQKFKFANLYITETDDE 115
 QY 130 KYLMK-DLTLDDQAYGAVENAKDIIACGPDINTFTFSLDLYMGSSGFYKVVYKIQKAV 188
 DB 116 KVMRNEPFLDQSRSAVYNDIIVGNPDKTIFQDTEYI---RMVYPTVYKAKTL 172
 QY 189 TENVKGIQFTSDSDICIGISFPALQAPSFNSPQIFRDRDTQCLIPCAIDQDPYR 248
 DB 173 TSEVATATGLDASSNIGLIFYPALQIAPL-----MFEKK---RCLIPAGIDQDPYR 222
 QY 249 MRDVAPRIGYPRPALLHSTFPALQAGTQKMSASDPNSIFLTDRAKQIKTVNKHAFS 308
 DB 223 LQRDIAESIGYKAAQIHSKPLPLGPGKSSSNPEATYIVDPKVERKIMKYAS 282
 QY 309 GGRDTEIHRQFGNCDVVSFMYLTFPLEDDD-KLEQIRKDYTSAMLTGELKKALIEV 367
 DB 283 GGQPTIELHRKYGNEIDVPPQMLYFFBEDNRIKEIEEYRSKMLTGLKQILIDK 342
 QY 368 LQPLIAEHQARRKVTDEIVKEFTPKIS 397
 DB 343 LNNFLEHR-RRREAKEHIVHFKYDKIA 371

RESULT 9
 SYM PYRAB STANDARD; PRT; 385 AA.
 AC Q9UY11;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TRPS).
 GN TRPS OR PAB111.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=29292;
 RN SEQUENCE FROM N.A.
 RC STRAIN=GES / Orsay;
 RA Heilig R.;
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure and evolution."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Trp) = AMP + dihydrophosphate + L-tryptophanyl-tRNA (Trp).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/annouce/> or send an email to license@isb-sib.ch).

DR EMBL; AJ248288; CAB50601.1; .
 DR InterPro; IPR002306; Trp-tRNA-synt-1b.
 DR InterPro; IPR002305; tRNA-synt-1b.
 DR InterPro; IPR001412; tRNA-synt-1.
 DR Pfam; PF00579; tRNA-synt-1b; 1.
 DR PRINTS; PR01039; TRNASYNTHTRP.
 DR TIGRFAMs; TIGR00233; trps; 1.
 DR PROSITE; PS00178; AA TRNA LIGASE I; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 FT SITE 82 90 "HIGH" REGION.
 FT SITE 253 257 "KMSKS" REGION.
 SQ SEQUENCE 385 AA; 45100 MW; 4C29D0144976B12 CRC64;
 Query Match 37.9%; Score 803; DB 1; Length 385;
 Best Local Similarity 45.3%; Pred. No. 1.2e-58;
 Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

QY 12 EDF-VDPMTVQTSSAKGIDYDYLIVRFSSSKIDKELINRIERATGQRPHPHPLRGIFFSH 70
 DB 3 EDFKTPMEVEG---VDYNKLIHFPGTSLTEELKTAETLTSPLFRKRFPSH 58
 QY 71 RDMQVLDAYENKKEFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVLVIQMTDDEK 130
 DB 59 RYDVKVLQDYEGRGPFYLYTGRGSG-PWHIGHIIPFATKMLQKFGVNLVIQITDDEK 117
 QY 131 YLMK-DLTLDDQAYGAVENAKDIIACGPDINTFTFSLDLYMGSSGFYKVVYKIQKAVT 189
 DB 118 FLFKENLTFEDTKHAWAYENIIDIIVGPDPTFTFQNSER---TKIYEMAPIAKKIN 173
 QY 190 FQVQKGIQFTSDSDICIGISFPALQAPSFNSPQIFRDRDTQCLIPCAIDQDPYRPM 249
 DB 174 FEMAAVAVGFTQSQSIGMIFFPALQIAPL-----PEKR---RCLIPALIDQDPYRML 223
 QY 250 TEDVAPRIGYPRPALLHSTFPALQAGTQKMSASDPNSIFLTDRAKQIKTVNKHAFSG 309
 DB 224 GRDFASISLGYYKTAIHSKFPVSLTSLGKMSASKPEATYIVLTDSPEDVEKVMFKALTG 283
 QY 310 GRDTEIHRQFGNCDVVSFMYLTFPLEDDDKLEQIRKDY---TSGAMLTGELKKALIE 366
 DB 284 GRPTLKEQREKGEPEKCVFWMLEIFFEEDK-KLERYACKNGELTGECCKRYLIS 341
 QY 367 VLQPLIAEHQARRKVTDEIVK 388
 DB 342 KIQEPLKEHQRRKKAKEQIEK 363

RESULT 10
 SYM PYRHO STANDARD; PRT; 386 AA.
 AC O59584;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TRPS).
 GN TRPS OR PH1921.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Hatakeyama Y., Hino Y., Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y., Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamizawa M., Ohfuku Y., Funahashi T., Tanaka T., Kudoh Y., Yamazaki U., Kushida N., Oguchi A., Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K., Masuchi Y., Shizuya H., Kikuchi H.;
 RA "Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";

RL DNA Res. 5:55-76(1998).
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Tyr) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA (Tyr).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in, no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AP000007; BA31046.1; ALT. INIT.
 DR InterPro: IPR002306; Trp_cRNA-synt_1b.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 DR PROSITE: PS00178; AA-TRNA-LIGASE_1; 1.
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KM Complete proteome.
 FT SITE 82 90 "HIGH" REGION.
 FT SITE 253 257 "KMSKS" REGION.
 SQ SEQUENCE 386 AA; 45305 MW; 9E3C392F4028B2DD CRC64;
 Query Match 37.6%; Score 796; DB 1; Length 386;
 Best Local Similarity 44.0%; Pred. No. 4.5e-58;
 Matches 172; Conservative 66; Mismatches 125; Indels 26; Gaps 9;
 QY 11 EEDFVDPWTVQTSAGKIDYDYLIVFGSSKIDKELINRIERATGQRPHHFLRGIFFSH 70
 DB 3 EEFKVTPEWVEGV---VDYDKLIKHFQTSPLTDLTKTAELTKSELPIFFRRKKFFFSH 58
 QY 71 RDMAOVLDAYENKKPFYLITGRGSSSEAMVGHILPIFTKMLQDVENVVLYQMTDDK 130
 DB 59 RVDVLILKDYEBEGGFELYGRGPGS-G-PMHIGHILPFPATKMLQDEKGVNLYIQITDDEK 117
 QY 131 YLWKD-LTLQAYGDAVENAKDIIACGPDINKTPIFSDLDYMGSSGFYKVVVKIOGHT 189
 DB 118 FLKRENTLPDTRKMAVDNLIDIIAVGFDPDKTIFPNSEF---TIIYEMALPIAKKIN 173
 QY 190 FNOVKIGFETDSDICIGISFPALQAPSFNSFPQIFRDRDIOCLIPCAIDODPYFRM 249
 DB 174 FSNAAKAVFGFTBQSKIGMIFFPALQINPFT-----FERK--RCILIPALIDODPYWRL 223
 QY 250 TRVVAIRIGYPPKALHSTFPFALQAGQTKASADPNSSIFLDTAKQITKYNKKAFAFG 309
 DB 224 QRFPAESLGYKTAALHSEKVPSPISLISGKMSASKEPETAIVLTDSPEDVEKTKWKFLLTG 283
 QY 310 GRDTIEHRQFGNCVDVDSFMYLTFPLEDDDKLEQIRKDY---TSGAMLTGELKKALIE 366
 DB 284 GRDTIEHRQFGNCVDVDSFMYLTFPLEDDDKLEQIRKDY---TSGAMLTGELKKALIE 366
 QY 367 VLOPLIAHQARRKEVTDEIVKEFMTPRKLS 397
 DB 342 KIQEFLKEHQRRRK-AEKLVEKFKYTKGKA 371
 RESULT 11
 ID SYM PYRAE STANDARD; PRT; 375 AA.
 AC 082TUS;
 DT 15-JUN-2002 (Rel. 41, Last Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan-tRNA ligase)
 DE (TrpRS).
 OS TRPS OR PAE3091.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Thermoproteaceae; Pyrobaculum.
 NCBI_TaxID=13773;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX PubMed=11792863;
 RA Fitz-Gibbon S.T., Lader H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 RT aerophilum";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Tyr) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA (Tyr).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE009911; AAL64664.1; -
 DR InterPro: IPR002306; Trp_cRNA-synt_1b.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR TIGRFAMs: TIGR00233; trps; 1.
 DR PROSITE: PS00178; AA-TRNA-LIGASE_1; 1.
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KM Complete proteome.
 FT SITE 81 89 "HIGH" REGION.
 FT SITE 258 262 "KMSKS" REGION.
 SQ SEQUENCE 375 AA; 43178 MW; 3DDBF85DA680F16 CRC64;
 Query Match 34.7%; Score 735; DB 1; Length 375;
 Best Local Similarity 41.6%; Pred. No. 4.5e-53;
 Matches 162; Conservative 79; Mismatches 130; Indels 18; Gaps 11;
 QY 11 EEDFVDPWTVQTSAGKIDYDYLIVFGSSKIDKELINRIERATGQRPHHFLRGIFFS 69
 DB 2 EEFVVTPEWVEGR---VDYDKLIKHFQTSPLTDLTKTAELTKSELPIFFRRKKFFFS 56
 QY 70 HRDMNOVLDAVENKKPFYLITGRGSSSEAMVGHILPIFTKMLQDVENVVLYQMTDDE 129
 DB 57 HRDFDPMKMGHGRPMALYTGGRGPGS-G-PVHIGHWPMWILLKWFSDKFGLEVYFQITDDE 115
 QY 130 K-YLWKD-LTLQAYGDAVENAKDIIACGPDINKTPIFSDLDYMGSSGFYKVVVKIOGHT 188
 DB 116 KFYDDEMKLEETWMAVENALDVIALGSPERHLIDTKOI--KPLYPVAVRAKKL 172
 QY 189 TFNOVGIGFETDSDICIGISFPALQAPSFNSFPQIFRDRDIOCLIPCAIDODPYFR 248
 DB 173 TMNTVATGTFDSTNIGLIFPSLOIAVAFPT--ELRRRATPV--LIPCAIDODPYFR 228
 QY 249 MTRDVAIRIGYPPKALHSTFPFALQAGQTKASADPNSSIFLDTAKQITKYNKKAFAFG 308
 DB 229 IARDIDALGYRPPSTLYSKFIALTG-ESKMSASNPDAIYLTDEKTVRRKV-WNAFT 286
 QY 309 GGRDTIEHRQFGNCVDVDSFMYLTFPLEDDDKLEQIRKDY---TSGAMLTGELKKALIEVL 368
 DB 287 GGRDTIEHRQFGNCVDVDSFMYLTFPLEDDDKLEQIRKDY---TSGAMLTGELKKALIEVL 368
 QY 369 QPLIAHQARRKEVTDEIVKEFMTPRKLS 397
 DB 347 TKFLKEHRRERKARSK-VDEYRLSVKLS 374
 RESULT 12
 ID SYM HALNI STANDARD; PRT; 380 AA.
 AC 09H86;
 DT 15-JUN-2002 (Rel. 41, Last Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)

15-JUN-2002 (Rel. 41, Last annotation update)
 DE TrypCophanY1-ERNA synthetase (EC 6.1.1.2) (TrypCophan--tRNA ligase)
 GN TRPS OR TRPS2 OR VNG2232G.
 OC Halobacterium sp. (strain NRC-1).
 OS Archaea; Euryarchaeota; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 NCBI_TaxId=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahaitas G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsen V., Shroog J.,
 RA Sattar S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,
 RA Lettner B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddock D.G., Jablonka P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isebaugh T.A., Peck R.F., Pohlman C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.,
 RA "Genome sequence of Halobacterium species NRC-1."
 RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 CC -1- CARBONYL ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE05109; AAC20355.1; -
 DR InterPro: IPR002306; Trp_tRNA-synt_1b.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRANSSYNTHTRP.
 DR TIGRFAMs: TIGR00233; ttps; 1.
 DR PROSITE: PS00178; AA TRNA LIGASE I; 1.
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 KM Complete proteome.
 FT SITE 74 / 82 "HIGH" REGION.
 FT SITE 249 / 253 "KMSK" REGION.
 SQ SEQUENCE 380 AA; 41936 MW; 0F7B2B955386404F CRC64;
 Query Match 25.3%; Score 534.5; DB 1; Length 380;
 Best Local Similarity 35.4%; Pred. No. 1.4e-36;
 Matches 141; Conservative 76; Mismatches 140; Indels 41; Gaps 17;
 QY 10 AEDFVDPWTQTSAGKIDYDKLIVRFSSKIDKELINIRATGQRPHHFLRGIFPS 69
 DB 3 AGDNVTPYAVASDD---LDYEKTLARFGADELTDQGRAPF-----DHPVNLGLFVA 53
 QY 70 HDNANQVLDAYENKPFYLYTGRGSEEMANGHLIFFTKMDQVENVPLVQMTDDE 129
 DB 54 GBDVDFLTAAGSOS---IVTGVGPGG-PMHGHANVVFARLDQEFARVYVPLSDSE 108
 QY 120 KYLMKDLTLDAQVDAVE--NAKDIACGFDINKTFIF---SLDADVA-GMSGSGYKVVVI 184
 DB 109 KYMFDQTPAET-GDYLRANLADLAVGDDPELTRIVUTRDAVDVYLPLATPAAGV--- 164
 QY 185 QKHTVFNOYKGIFFGDSICKISFPALQAAPEFSNSFPQIRDRDTDIQCLIPCAIDP 244
 DB 165 -RHATLQNYG-----EPDNVQAFYPAVQTAHL---LPQLVHG--EHETLVIPLAVDD 213
 QY 245 PYFRMTRDVAIRIGV--KEALLHSTFPALQAGQKMSADPNSGIFLTDRAKOIKTV 302
 DB 214 PYHRSRDVAARARFVVGKFGALMQLFSLAG-PEKMS--AGVIRLTDSPDVTREYV 271
 QY 303 NGAHSGGKDTTEHRQFGKNCDDVDSFMVLTFFLEDD-KLEQIRKDYTSAGMLTGEIK 361

272 RHATYTGASVSEHRAAGVPAEDVPPQYLSAFEPDPAELARLEREYRAGDLSGLK 331
 QY 362 KALIEVLOPLIAHQARKREVTDEIVKEPMTRKLSFD 399
 DB 332 DLAADRITFEFLAAHQRRRAALGD--VTEALDAFRILTDD 367
 RESULT 13
 ID SYM METUA STANDARD; PRT; 370 AA.
 AC 058810;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE TrypCophanY1-ERNA synthetase (EC 6.1.1.2) (TrypCophan--tRNA ligase)
 GN TRPS OR MJ1415.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 NCBI_TaxId=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8668087;
 RA Bult C.J., White O., Olsen G.D., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kierliavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weissrock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fudmann J.L., Nguyen D.,
 RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadov P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurest M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Frazer C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.";
 RT Science 273:1058-1073 (1996).
 CC -1- CARBONYL ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U67582; AAB99425.1; -
 DR TIGR: MJ1415; -
 DR InterPro: IPR002306; Trp_tRNA-synt_1b.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRANSSYNTHTRP.
 DR TIGRFAMs: TIGR00233; ttps; 1.
 DR PROSITE: PS00178; AA TRNA LIGASE I; FALSE NEG.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KM Complete proteome.
 FT SITE 75 / 83 "HIGH" REGION.
 FT SITE 255 / 259 "KMSK" REGION.
 SQ SEQUENCE 370 AA; 42660 MW; E6C71107C82B59D CRC64;
 Query Match 19.4%; Score 409.5; DB 1; Length 370;
 Best Local Similarity 30.9%; Pred. No. 2.6e-26;
 Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;
 QY 17 PWTQTSAGKIDYDKLIVRFSSKIDKELINIRATGQRPHHFLRGIFPSHDNANQV 76
 DB 8 PW--ETPAV--IDYKTKMEFGVKKPIVDVLDLKEE-----HHFRNNIILIGHDRPFR 57

QY 77 LDAYENKKPFLYLYTGRGSSSEAMHVGHLPIFIKMLQ---DVFNVLVIQMTDDEKYLK 133
 DB 58 VDAIKNNKEBAVVSNGMPSGK-MHFGKXWVDDLLKFKQXTNINIP1-----ADLEAYMA 112
 QY 134 KDITLDQAYGDAV-ENAKDIIACGFPDINKTFTISDDLDMGMSGGFYKNVVK1-QKHVTEN 191
 DB 113 RNNSPFTYKELALNEVITNYIALGLDPEKINVLQSKYQKV-----KDLILISKKTMS 167
 QY 192 QVAGIGFTDSDICIGKISFPALQAPSFNSPQIFNDRD--DIQCLIPCAIDODPYFRM 249
 DB 168 EMATIGFKETNIGHVFADIVVADL---HPOLDENLSPEKPVVVPVIGIDDPHIL 224
 QY 250 TRDAVR---IGYKPALHSTFEPALQAGTQMSASDPNSSIFLDTAKQIKTKYKKA 306
 DB 225 TRDIARAKFKFTIPPSSTYHRFMVTLGG--KMSSEKPEFALFLDDEKTVKKIIFS-A 281
 QY 307 FSGGRTIEHROFGG---NCDVDSFMYLTFFLEDDDKLEQIRKDYTSAMLTGELKKA 363
 DB 282 KTGGRSTLEHKKYGGVPECVVYELFLY--HLILDKELAETIYQKCRSGELTGCKCKM 339
 QY 364 LIEVLQPLIAHQARKKVTDEIVK 388
 DB 340 AVERVVEFLKDLKKEQAKELAVK 364

RESULT 14

SYN_METHH STANDARD; PRT; 364 AA.

AC 026352;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TrpS)
 GN TRPS OR MTH251.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H; /
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pochier B., Qiu D.,
 RA Spadefora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Patelwell D., Fradnakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RA "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genome.";
 RT J. Bacteriol. 179:7135-7155(1997).
 RL J. CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Trp) = AMP +
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophanyl-tRNA (Trp).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AE00812; AAB84757.1; -
 DR InterPro; IPR002306; Trp tRNA-synt_1b.
 DR InterPro; IPR002305; tRNA-synt_1b.
 DR InterPro; IPR00412; tRNA-synt_1.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PRO1039; TRNASYN1TRP.
 DR TIGRfam; TIGR00233; trps; 1.

DR PROSITE; PS00178; AA tRNA LIGASE I; FALSE NEG.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 FT SITE 70 78 "HIGH" REGION.
 FT SITE 251 255 "KMSKS" REGION.
 SQ SEQUENCE 364 AA; 41301 MW; C2F348903338F61D CRC64;
 Query Match 18.8%; Score 397.5; DB 1; Length 364;
 Best local similarity 27.9%; Pred. No. 2.5e-25;
 Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;

QY 15 VDPWVQTSANGIDYDKLIVRFGSSKIDKELINRIERTGQRPHHFLRGIFFSRDKN 74
 DB 2 IDPW-----GSAR-LIEVODLIENFVGRPF-SEVLDEV-----DEPSWLMRGIIIFGRDVE 50
 QY 75 QVLDAYENKKPFLYLYTGRGSSSEAMHVGHLPIFIKMLQDVFNVLVIQMTDDEKYLK 134
 DB 51 RIISAKKGEDPAVVTGMPSGR-MHFGKXWVDDLLKFKQXTNINIP1-----ADLEAYMA 108
 QY 135 DLTLDAQYDAVEN-AKDIACGFPDINK-----TFISDDLDMGMSGGFYKNVVK 183
 DB 109 GVDFEDSRRIALIEVYIAGYIALGLDEKONHIVYLQSENLMWEDLAYV----- 156
 QY 184 IQGHTFNCVKGIFGFTSDICIGKISFPALQAPSFNSPQIFNDRDITQCLIPCAIDQ 243
 DB 157 LAGKVNFEELRAIYGFSTGSMAMVAPLIQVSDILHPOLDLEGPR---PVIYPGPQ 213
 QY 244 DPEFMTTRVAPRI---GYKPALHSTFEPALQAGTQMSASDPNSSIFLDTAKQIK 299
 DB 214 DPHIRLRIIDIAFRDRYFIIIPSSYHFMGLTGG--KMSNRPKAIFFISDTPFEAE 271
 QY 300 TKVKNKAFSGDRTIEHROFGNCDVDSFMYLTFFLE-DDDKLEQIRKDYTSAMLTG 358
 DB 272 AKI-RNAKTGGRSTLEHKKYGGVPECVVYELFLYLYHMGSDSRLEIYESCRNGLMCG 330
 QY 359 ELKKALIEVLQPLIAHQARKK 381
 DB 331 ECKNNTAEIRIRKFEELSVKREK 353

RESULT 15

SYN_ENCCU STANDARD; PRT; 134 AA.

AC 096771;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TrpS) (Fragment).
 GN Encephalitozoon cuniculi.
 OS Encephalitozoon cuniculi.
 OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
 OX NCBI_TaxID=6035;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277683; PubMed=9615449;
 RA Peyretalade B., Brousseau V., Peyret P., Metenier G., Gouy M.,
 RA Vivares C.P.;

RT "Microsporidia, mitochondrial protists, possess a 70-kDa heat shock
 RT protein gene of mitochondrial evolutionary origin.";
 RL Mol. Biol. Evol. 15:683-689(1998).
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA (Trp).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:28:00 ; Search time 15.5539 Seconds
(without alignments)
2478.466 Million cell updates/sec

Title: US-09-813-718-10_COPY_71_471

Perfect score: 2116
Sequence: 1 SNHGDATEAEEDFVDPMTV.....VTDEIVKEFMPDKLSDFDQ 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2101	99.3	471	1 A41706	cryptophan-tRNA 11
2	2028.5	95.9	475	1 YWBO	cryptophan-tRNA 11
3	1946.5	92.0	475	1 YWRBPR	cryptophan-tRNA 11
4	1938	91.6	481	2 S50053	cryptophan-tRNA 11
5	1210	57.2	395	2 S58157	hypothetical prote
6	1163	55.0	432	2 S51901	cryptophan-tRNA 11
7	907	42.9	386	2 C90190	cryptophan-tRNA 11
8	803	37.9	385	2 C75020	cryptophan-tRNA 11
9	626.5	29.6	301	2 G71206	cryptophan-tRNA 11
10	534.5	25.3	380	2 G84373	cryptophan-tRNA 11
11	409.5	19.4	370	2 F64476	cryptophan-tRNA 11
12	397.5	18.8	364	2 E69131	cryptophan-tRNA 11
13	386	18.2	134	2 T43806	cryptophan-tRNA 11
14	370.5	17.5	420	2 E69461	cryptophan-tRNA 11
15	353	16.7	374	2 D72477	probable tyrosyl-c
16	269.5	12.7	513	2 F84371	cryptophan-tRNA 11
17	192	9.1	323	2 H69346	tyrosyl-tRNA synth
18	188	8.9	364	2 E72512	probable tyrosyl-t
19	185.5	8.8	341	2 D95260	cryptophan-tRNA 11
20	185.5	8.8	341	2 G98125	cryptophan-tRNA 11
21	174.5	8.2	341	2 B86633	cryptophan-tRNA 11
22	172	8.1	394	2 A45999	tyrosine-tRNA 11ga
23	170.5	8.1	366	2 S75410	tyrosine-tRNA 11ga
24	169	8.0	395	2 H70385	cryptophan-tRNA 11
25	167	7.9	351	2 E75438	cryptophan-tRNA 11
26	166.5	7.9	346	2 B71496	cryptophan-tRNA 11
27	164.5	7.8	895	2 A86410	protein F3M18.22
28	163.5	7.7	353	2 E70100	cryptophan-tRNA 11
29	163	7.7	460	2 C84750	probable tyrosyl-t

30	159.5	7.5	337	2 F71300	cryptophan-tRNA 11
31	158.5	7.5	375	2 B75072	tyrosyl-tRNA synth
32	158	7.5	346	2 C81654	cryptophan-tRNA 11
33	154.5	7.3	102	2 T44994	probable tyrosyl-t
34	152	7.2	344	2 H86590	cryptophan-tRNA 11
35	152	7.2	344	2 C72034	cryptophan-tRNA 11
36	149.5	7.1	327	2 C84374	tyrosyl-tRNA synth
37	149.5	7.1	408	2 T03741	probable tyrosine-
38	146	6.9	339	2 E64676	cryptophan-tRNA 11
39	146	6.9	375	2 F71093	tyrosine-tRNA 11ga
40	145	6.9	337	2 A11066	cryptophan-tRNA 11
41	138	6.5	319	2 H69102	tyrosine-tRNA 11ga
42	135.5	6.4	328	2 C72370	cryptophan-tRNA 11
43	135	6.4	365	2 E82052	cryptophan-tRNA 11
44	132.5	6.3	343	2 S73024	cryptophan-tRNA 11
45	132	6.2	328	1 YWBSF	cryptophan-tRNA 11

ALIGNMENTS

RESULT 1
A41706
cryptophan-tRNA ligase (EC 6.1.1.2) [similarity] - human
N/Alternate names: interferon-inducible protein IFP53; peptide-chain release factor homo
C/Species: Homo sapiens (man)
C/Date: 19-May-2000 #sequence revision 19-May-2000 #text change 03-Jun-2002
C/Accession: A41633; A41706; S19246; UN0676; JH0533; S26287
R/Flecker, J.; Rasmussen, H.H.; Juellesen, J.
Proc. Natl. Acad. Sci. U.S.A. 88, 11520-11524, 1991
A/Title: Human interferon gamma potentially induces the synthesis of a 55-kDa protein (gamma
A/Reference number: A41633; MUID:92107982; PMID:1763065
A/Accession: A41633
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-471 <RUB>
A/Cross-references: GB:X59892; NID:G30820; PIDN:CAA42545.1; PID:G30821
R/Rubin, B.Y.; Anderson, S.L.; Xing, L.; Powell, R.J.; Tate, W.P.
J. Biol. Chem. 266, 2445-24248, 1991
A/Title: Interferon induces cryptophan-tRNA synthetase expression in human fibroblasts.
A/Reference number: A41706; MUID:92105071; PMID:1761529
A/Accession: A41706
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-471 <RUB>
A/Cross-references: GB:M77804; NID:G184656; PIDN:AAA67324.1; PID:G184657
R/Buwlit, U.; Flohr, T.; Boettger, E.C.
EMBO J. 11, 489-496, 1992
A/Title: Molecular cloning and characterization of an interferon induced human cDNA with
A/Reference number: S19246; MUID:92164636; PMID:1537332
A/Accession: S19246
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-423, 'R', 425-471 <BUW>
A/Cross-references: EMBL:X62570; NID:G32708; PIDN:CAA44450.1; PID:G32709
R/Frolova, L.Y.; Grigorjeva, A.Y.; Sudomoina, M.A.; Kiselev, L.L.
Gene 128, 237-245, 1993
A/Title: The human gene encoding cryptophan-tRNA synthetase: Interferon-response element
A/Reference number: JN0676; MUID:9329292; PMID:7685728
A/Accession: JN0676
A/Molecule type: DNA
A/Residues: 1-141, 182-471 <FROL>
A/Cross-references: GB:X67918; GB:S62837; NID:G37968; GB:X67919; NID:G37969; GB:X67920; F
4; NID:G37974; GB:X67925; GB:S62855; NID:G37975; GB:X67926; GB:S62856; NID:G37976; GB:X6
A/Note: the authors translated the codon GGG for residue 55 as Cys and GAG for residue 34
R/Frolova, L.Y.; Sudomoina, M.A.; Grigorjeva, A.Y.; Zinovjeva, O.L.; Kiselev, L.L.
Gene 109, 291-296, 1991
A/Title: Cloning and nucleotide sequence of the structural gene encoding for human trypt
A/Reference number: JH0533; MUID:92112058; PMID:1765274
A/Accession: JH0533
A/Molecule type: mRNA

A:Residues: 1-212, 'GD', 215-471 <FRO2>
 A:Cross-references: GB:M61715; NID:g340367; PIDN:AAA61298.1; PID:g340368
 A:Experimental source: fibroblast
 C:Genetics:
 A:Gene: GDB:WARS; IEP53
 A:Cross-references: GDB:119632; OMIM:191050
 A:Map position: 14q23-14q31
 A:Intons: 33/3; 105/1; 141/2; 181/2; 242/2; 276/1; 313/3; 371/3; 418/3
 C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
 C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis
 F:19-64/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 99.3%; Score 2101; DB 1; Length 471;
 Best Local Similarity 99.5%; Pred. No. 2e-163;
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNHGPDTEAEEDFVDPWTVQTSAGKIDYDKLIVFGSSKIDKELINRIERATGQRP 60
 DB 71 SNHGPDTEAEEDFVDPWTVQTSAGKIDYDKLIVFGSSKIDKELINRIERATGQRP 130
 QY 61 FLRGGIFFSHRDNNQVLDAVENKKPFYLYTGRGSSSEAMVGHILPFIETKMLQDVNVP 120
 DB 131 FLRGGIFFSHRDNNQVLDAVENKKPFYLYTGRGSSSEAMVGHILPFIETKMLQDVNVP 190
 QY 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKN 180
 DB 191 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKN 250
 QY 181 VVKIQKHTFNQVKGIFGFTSDSDICGISPPALQAPSPNSFPQIFRRTDIQCLIPCA 240
 DB 251 VVKIQKHTFNQVKGIFGFTSDSDICGISPPALQAPSPNSFPQIFRRTDIQCLIPCA 310
 QY 241 IDDDPYFRMTRDVAPRIGYKPPALHSTFPALQAGTQMSADPNSSIFLDTAKQIKT 300
 DB 311 IDDDPYFRMTRDVAPRIGYKPPALHSTFPALQAGTQMSADPNSSIFLDTAKQIKT 370
 QY 301 KVNKAHFGSGRDTIEHRQFGNCNDVVSFMYLTFLEDDDKLEQIRKDYTSAMLTGEL 360
 DB 371 KVNKAHFGSGRDTIEHRQFGNCNDVVSFMYLTFLEDDDKLEQIRKDYTSAMLTGEL 430
 QY 361 KKAILEVLOPLIAEHQARRKEVTDEIVKEFMPRKLSFDFQ 401
 DB 431 KKAILEVLOPLIAEHQARRKEVTDEIVKEFMPRKLSFDFQ 471

RESULT 2
 YRBO
 tryptophan-tRNA ligase (EC 6.1.1.2) [validated] - bovine
 N:Alternate names: tryptophanyl-tRNA synthetase
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 30-Sep-1992 #sequence revision 30-Sep-1992 #text_change 03-Jun-2002
 C:Accession: A40279; JN0354; S10460; S14540
 R:Garret, M.; Rajoc, B.; Trezeguet, V.; Labouesse, J.; Merle, M.; Gandr, J.C.; Benedet
 Biochemistry 30, 7809-7817, 1991
 A:Title: A mammalian tryptophanyl-tRNA synthetase shows little homology to prokaryotic
 A:Reference number: A40279; MUID:91329348; PMID:1907847
 A:Accession: A40279
 A:Molecule type: mRNA
 A:Residues: 1-475 <GAR>
 A:Cross-references: GB:M74074; EMBL:X53918; NID:g163798; PIDN:AAA30799.1; PID:g163799
 A:Experimental source: pancreas
 A:Note: the authors translated the codon CTG for residue 347 as Ala and CAG for residue
 R:Zargarova, T.A.; Kovalova, G.K.; Favorova, O.O.; Aevina, H.B.; Telezhinskaya, I.H.
 Biorg. Khim. 15, 1307-1311, 1989
 A:Title: Amino acid sequence of several peptides of tryptophanyl-tRNA synthetase from ca
 A:Reference number: JN0354; MUID:90211408; PMID:2631684
 A:Accession: JN0354
 A:Molecule type: protein
 A:Residues: 112-124;282-287, 'N', 288, 'F', 289-292, 'Q', 293-294, 'IR', 336-353;423-441,443-445
 A:Experimental source: liver
 A:Note: this paper is in Russian
 C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis
 F:24-69/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 95.9%; Score 2028.5; DB 1; Length 475;
 Best Local Similarity 95.8%; Pred. No. 1.7e-157;
 Matches 384; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 1 SNHGPDTEAEEDFVDPWTVQTSAGKIDYDKLIVFGSSKIDKELINRIERATGQRP 60
 DB 76 SGEGLATEAEEDFVDPWTVQTSAGKIDYDKLIVFGSSKIDKELINRIERATGQRP 135
 QY 61 FLRGGIFFSHRDNNQVLDAVENKKPFYLYTGRGSSSEAMVGHILPFIETKMLQDVNVP 120
 DB 136 FLRGGIFFSHRDNNQVLDAVENKKPFYLYTGRGSSSEAMVGHILPFIETKMLQDVNVP 195
 QY 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKN 180
 DB 196 LVIQMTDEKYLKMDLTLDQAYGDAVENAKD-IIICGPDINKTFIFSDLDYMGSSGFYKN 254
 QY 181 VVKIQKHTFNQVKGIFGFTSDSDICGISPPALQAPSPNSFPQIFRRTDIQCLIPCA 240
 DB 255 VVKIQKHTFNQVKGIFGFTSDSDICGISPPALQAPSPNSFPQIFRRTDIQCLIPCA 314
 QY 241 IDDDPYFRMTRDVAPRIGYKPPALHSTFPALQAGTQMSADPNSSIFLDTAKQIKT 300
 DB 315 IDDDPYFRMTRDVAPRIGYKPPALHSTFPALQAGTQMSADPNSSIFLDTAKQIKT 374
 QY 301 KVNKAHFGSGRDTIEHRQFGNCNDVVSFMYLTFLEDDDKLEQIRKDYTSAMLTGEL 360
 DB 375 KVNKAHFGSGRDTIEHRQFGNCNDVVSFMYLTFLEDDDKLEQIRKDYTSAMLTGEL 434
 QY 361 KKAILEVLOPLIAEHQARRKEVTDEIVKEFMPRKLSFDFQ 401
 DB 435 KKAILEVLOPLIAEHQARRKEVTDEIVKEFMPRKLSFDFQ 475

RESULT 3
 YRBP
 tryptophan-tRNA ligase (EC 6.1.1.2) [validated] - rabbit
 N:Alternate names: tryptophanyl-tRNA synthetase
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C>Date: 30-Sep-1992 #sequence revision 13-Feb-1998 #text_change 03-Jun-2002
 C:Accession: A35904; S37396
 R:Lee, C.C.; Craigden, W.J.; Muzny, D.M.; Harlow, E.; Caskey, C.T.
 Proc. Natl. Acad. Sci. U.S.A. 87, 3508-3512, 1990
 A:Title: Cloning and expression of a mammalian peptide chain release factor with sequence
 A:Reference number: A35904; MUID:90239043; PMID:2185472
 A:Accession: A35904
 A:Molecule type: mRNA
 A:Residues: 1-475 <LEB>
 A:Cross-references: GB:M33460
 R:Polova, L.Y.; Dalphin, M.E.; Justesen, J.; Powell, R.J.; Druegan, G.; McCaughan, K.K.;
 EMBO J. 12, 4013-4019, 1993
 A:Title: Mammalian polypeptide chain release factor and tryptophanyl-tRNA synthetase are
 A:Reference number: S37396; MUID:94009008; PMID:8404867
 A:Accession: S37396
 A:Molecule type: mRNA
 A:Residues: 166-177 <FRO>
 C:Genetics:
 A:Gene: WRS
 C:Complex: homodimer [validated, MUID:94009008]
 C:Function:
 A:Description: EC 6.1.1.2 [validated, MUID:94009008]; catalyzes the ATP-dependent format
 A:Note: mammalian WRS (tryptophanyl-tRNA synthetase) and eRF (polypeptide chain release
 C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
 C:Keywords: aminoacyl-tRNA synthetase; ATP; homodimer; ligase; metalloprotein; protein bi
 F:23-68/Domain: amino acid-tRNA ligase repeat homology <ATL>
 F:174-177/Region: ATP-binding motif (HXGH)

Query Match 92.0%; Score 1946.5; DB 1; Length 475;
 Best Local Similarity 90.8%; Pred. No. 8.2e-151;
 Matches 364; Conservative 21; Mismatches 15; Indels 1; Gaps 1;

QY 2 NHG-PDATEABEDFVDPWTVQTSAGKIDYDKLIVRGSSKIDKELINRERATGQRPHP 60
 Db SHDDPEAVDDKDFVDPWTVQTSAGKIDYDKLIVGSSKIDKELINRERATGQRPHP 134
 QY 61 FLRRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVNVNP 120
 Db FLRRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVNVNP 134
 QY 121 LVQMTDDEKYLWMDLTLDQAYGDAVENADIIACGPDINKTFIFSLDYMGMSSGFFKYN 180
 Db LVQMTDDEKYLWMDLTLDQAYGDAVENADIIACGPDINKTFIFSLDYMGMSSGFFKYN 254
 QY 181 VKIQKHVTVNOVKGIFGFTSDICIGKISFPALQAPSPNSFPQIFRDRTDIOCLIPCA 240
 Db VKIQKHVTVNOVKGIFGFTSDICIGKISFPALQAPSPNSFPQIFRDRTDIOCLIPCA 314
 QY 241 IDDDPYFRMTRDVAPRIGYKRPALHSTFPALQAGATKMSADPNSSIFLTDAKQIKT 300
 Db IDDDPYFRMTRDVAPRIGYKRPALHSTFPALQAGATKMSADPNSSIFLTDAKQIKT 374
 QY 301 KVKHAFSGGRDITIEHRQFGNCVDVSPMYLTFFLEDDEKLEQIRKDYSSGAMLTGEL 360
 Db KVKHAFSGGRDITIEHRQFGNCVDVSPMYLTFFLEDDEKLEQIRKDYSSGAMLTGEL 434
 QY 361 KKALIEVLQPLIAEHQARRKVTDEIVKEFMTPRKLSFDPQ 401
 Db KKALIEVLQPLIAEHQARRKVTDEIVKEFMTPRKLSFDPQ 475

RESULT 4

550053
 tryptophan-tRNA ligase (EC 6.1.1.2) alpha-2 chain - mouse

C/Species: Mus musculus (house mouse)
 C/Date: 07-May-1995 #sequence_revision 10-Nov-1995 #text_change 03-Jun-2002

C/Accession: S50053; S50052; I49391; S31461; S31462
 R/Pajot, B.; Sarger, C.; Bonnet, J.; Garret, M.

J. Mol. Biol. 242, 599-603, 1994

A>Title: An alternative splicing modifies the C-terminal end of tryptophanyl-tRNA synthetase
 A/Reference number: S50052; MUID:95018226; PMID:7932716

A/Accession: S50053

A/Molecule type: mRNA
 A/Residues: 1-481 <PAJ>

A/Cross-references: EMBL:X69657; NID:G55437; PIDN:CAA49348.1; PID:G55438

A/Genetic: LSP
 A/Note: Intron position was determined by sequencing of genomic DNA

A/Accession: S50052

A/Molecule type: mRNA
 A/Residues: 1-475 <PAM>

A/Cross-references: EMBL:X69656; NID:G55435; PIDN:CAA49347.1; PID:G55436

A/Genetic: SSF
 R/Kiesel, L.L.

Biochimie 75, 1027-1039, 1993

A>Title: Mammalian tryptophanyl-tRNA synthetases.
 A/Reference number: I49391; MUID:94257729; PMID:7515282

A/Accession: I49391

A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA

A/Residues: 1-481 <RBS>

A/Cross-references: EMBL:X69657; NID:G55437; PIDN:CAA49348.1; PID:G55438

A/Genetic: <LSF>

A/Introns: 475/2
 A/Note: the list of introns may be incomplete; clone W13

C/Genetic: <SSP>

A/Genetic: WRS

A/Note: clone S5

C/Suprafamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
 C/Keywords: alternative splicing; aminoacyl-tRNA synthetase; ligase; protein biosynthesis
 F.1-481/Product: tryptophan-tRNA ligase alpha-2 chain long splice form #status predicted
 F.1-475/Product: tryptophan-tRNA ligase alpha-2 chain short splice form #status predicted
 F.1-475/Product: tryptophan-tRNA ligase repeat homology <ATL>

Query Match 91.6%; Score 1938; DB 2; Length 481;

Best Local Similarity 90.8%; Pred. No. 4, 1e-150;
 Matches 363; Conservative 21; Mismatches 16; Indels 0; Gaps 0;
 QY 2 NHG-PDATEABEDFVDPWTVQTSAGKIDYDKLIVRGSSKIDKELINRERATGQRPHP 61
 Db NCDSDATKASEDFVDPWTVQTSAGKIDYDKLIVGSSKIDKELINRERATGQRPHP 135
 QY 62 FLRRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVNVNP 121
 Db FLRRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVNVNP 195
 QY 122 LVQMTDDEKYLWMDLTLDQAYGDAVENADIIACGPDINKTFIFSLDYMGMSSGFFKYN 181
 Db LVQMTDDEKYLWMDLTLDQAYGDAVENADIIACGPDINKTFIFSLDYMGMSSGFFKYN 255
 QY 182 VKIQKHVTVNOVKGIFGFTSDICIGKISFPALQAPSPNSFPQIFRDRTDIOCLIPCA 241
 Db VKIQKHVTVNOVKGIFGFTSDICIGKISFPALQAPSPNSFPQIFRDRTDIOCLIPCA 315
 QY 242 IDDDPYFRMTRDVAPRIGYKRPALHSTFPALQAGATKMSADPNSSIFLTDAKQIKT 301
 Db IDDDPYFRMTRDVAPRIGYKRPALHSTFPALQAGATKMSADPNSSIFLTDAKQIKT 375
 QY 302 VNHAFSGGRDITIEHRQFGNCVDVSPMYLTFFLEDDEKLEQIRKDYSSGAMLTGEL 361
 Db VNHAFSGGRDITIEHRQFGNCVDVSPMYLTFFLEDDEKLEQIRKDYSSGAMLTGEL 435
 QY 362 KKALIEVLQPLIAEHQARRKVTDEIVKEFMTPRKLSFDPQ 401
 Db KKALIEVLQPLIAEHQARRKVTDEIVKEFMTPRKLSFDPQ 475

RESULT 5

558157
 hypothetical protein SPAC2F7.13c - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe
 C/Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 10-Dec-1999

C/Accession: S58157; T38561
 R/Gentles, S.; Church, C.M.

submitted to the EMBL Data Library, July 1995

A/Accession: S58157

A/Molecule type: DNA
 A/Status: preliminary

A/Residues: 1-395 <GEN>

A/Cross-references: EMBL:Z50142; NID:G1052783; PIDN:CAA90500.1; PID:G1052796

R/Gentles, S.; Church, C.M.; Barrett, B.G.; Rajandream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, July 1995

A/Reference number: Z21799

A/Accession: T38561
 A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-395 <GEN>

A/Cross-references: EMBL:Z50142; PIDN:CAA90500.1; GSPDB:GN00066; SPDB:SPAC2F7.13c

C/Genetic: SPDB:SPAC2F7.13c

A/Map position: 1
 C/Suprafamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 57.2%; Score 1210; DB 2; Length 395;
 Best Local Similarity 59.6%; Pred. No. 7, 2e-91;
 Matches 225; Conservative 60; Mismatches 91; Indels 8; Gaps 4;

QY 11 EEDFVDPWTVQTS---SAKGIDYDKLIVRGSSKIDKELINRERATGQRPHP 66
 Db EEDFVDPWTVQTS---SAKGIDYDKLIVRGSSKIDKELINRERATGQRPHP 63
 QY 67 FSHRDMNOVLDAENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVNVPLVQMT 126
 Db FSHRDMNOVLDAENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVNVPLVQMT 123
 QY 127 DDEKYLWMDLTLDQAYGDAVENADIIACGPDINKTFIFSLDYMGMSSGFFKYN 185

Db 124 DDEKFLPKQGVSLDCCORFARENAKOIIANGFDPKTFIFMNSYTV--GAFQNVRIA 161
QY 186 KHTVFNQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDIOCLIPCAIDDP 245
Db 182 KCTTANQSKACFGFTDSDISGKIHFAISIOAAPSFSFPHIFNGAKDIPCLIPCAIDDP 241
QY 246 YFMTDVAARIGYPRKALHSTFFPALQAQOTKMSADPNSSIFLDTAKQIKTKYKH 305
Db 242 YFRLTRVSGRLKFKKALHSTFFPALQAQOTKMSADPNSSIFLDTAKQIKTKYKH 301
QY 306 AFGSGRDTIEHRQFGNCVDVVSFMYLTFLEDDDLKQIRKDYTSGAMLTGELKKALI 365
Db 302 AFGSGGATIEHRKGNPNVDVAYQYLSFFLDDDLKQIKYNTYKAGTISTGEMGECI 361
QY 366 EVLQPLIAEHQARRKEVTDEIVKEFMT-PRKLSF 398
Db 362 KLLQGFVSDFOAARSKVDEATLDMFMDGSRKLEW 395

RESULT 6

S51901
tryptophan-tRNA ligase (EC 6.1.1.2) [similarity] - Yeast (*Saccharomyces cerevisiae*)

N/Alternate names: protein HR432; protein 00792; tryptophanyl-tRNA synthetase
C/Species: *Saccharomyces cerevisiae*
C/Date: 05-May-1995 #sequence_revision 03-Aug-1995 #text_change 03-Jun-2002

C/Accession: S51901, S59177; S66793
R/Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.
submitted to the EMBL Data Library, January 1995

A/Description: Sequence analysis of a 44kb DNA fragment of yeast chromosome XV including and a Delta.
A/Reference number: S51848
A/Accession: S51901

A/Molecule type: DNA
A/Residues: 1-432 <YAN>

A/Cross-references: EMBL:Z48149; NID:G663234; PIDN:CAA88164.1; PID:G663256
R/Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.
Yeast 11, 1069-1075, 1995

A/Title: Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV including the a delta element.
A/Reference number: S59156; MUID:96076631; PMID:7502582

A/Accession: S59177
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA

A/Residues: 1-432 <YAN>
A/Cross-references: EMBL:Z48149; NID:G663234; PIDN:CAA88164.1; PID:G663256
A/Note: the nucleotide/sequence was submitted to the EMBL Data Library, January 1995

R/Durand, P.; Hilger, F.; Portetelle, D.; Vandenbol, M.
submitted to the Protein Sequence Database, July 1996
A/Reference number: S66791

A/Accession: S66793
A/Molecule type: DNA

A/Residues: 1-432 <DUR>
A/Cross-references: EMBL:Z74839; NID:G1419947; PIDN:CAA99110.1; PID:G1419948; GSPDB:GN000
A/Experimental source: strain S288C
C/Genetics:

A/Gene: SGD:WRS1, WRS1, MIPS:YOL097C
A/Cross-references: SGD:S0005457
A/Map position: 15L

C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
C/Keywords: ligase
F:117-120/Region: ATP-binding motif (HXHG)

Query Match

Best local similarity 55.0%; Score 1163; DB 2; Length 432;
Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

QY 7 ATEAEDFVDPWVY-----QTSAGKIDYDCLIVRFGSSKIDKELINRIRATGQRPHH 61
Db 19 STDVKEQVTPWDEGGVDSGRQNIIDYDCLIKQFSTKPNSETLTKRFGVGRPEPHF 78
QY 62 LRGIIFPSHDMNOVLDAENKKPFYLYTGRGSSSEAMHVGHLPIFLTKMLQDVFNVPL 121

Db 79 LRKGLPFSSRDPFTKILDYEQGKPFYLYTGRGSSSDSMELGHMIPFVFTKKLQEVDPVL 138
QY 122 VIQMTDDEKYLK-DLTLDQAGDAVENAKDIIACGFDINKFTIESDLDYMGSSGFYKA 180
Db 139 VIELTDEKFLFGKRLTINDVKNFAENAKDIIAVFDPKNTFIFSDLYMG--GAFYET 196
QY 181 VVKIQHVTFNQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 240
Db 197 VVAVSRQITGTRKAVFGGRSDCIKHFASIQITAPSSFPNVLGPDKTPCLIPCA 256
QY 241 IDDPYFRMTDVAARIGYPRKALHSTFFPALQAQOTKMSADPNSSIFLDTAKQIKT 300
Db 257 IDDPYFRVCRDVAADKLTKSKPALHSTFFPALQOSTTMSADPNSSIFLDTAKQIK 316
QY 301 KYNKAFSGGRDTIEHRQFGNCVDVVSFMYLTFLEDDDLKQIRKDYTSGAMLTGEL 360
Db 317 KINKVAFSGQVADLHRELGPNVDVAYQYLSFFKDDVFLKECYDKYKSGELLGEM 376
QY 361 KKAILEVQPLIAEHQARRKEVTDEIVKEFMT-PRKL 396
Db 377 KKLCTETLOEFVAKAFQERRAQVDEETLDKFMVPHKL 412

RESULT 7

C90190
tryptophanyl-tRNA synthetase (trps) [imported] - *Sulfolobus solfataricus*

C/Species: *Sulfolobus solfataricus*
C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001

C/Accession: C90190
R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-Y
Jong, I.; Jeffries, A.C.; Kozier, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
submitted to Genbank, April 2001

A/Description: *Sulfolobus solfataricus* complete genome.
A/Reference number: A99139
A/Accession: C90190

A/Status: preliminary
A/Molecule type: DNA

A/Residues: 1-386 <KUR>
A/Cross-references: GB:AE006641; NID:G13813608; PIDN:AAK40778.1; GSPDB:GN00155
C/Genetics:

A/Gene: trps
C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 42.9%; Score 907; DB 2; Length 386;

Best local similarity 48.6%; Pred. No. 3, 66-87;
Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

QY 8 TEAEDFVDPWVYQVTSAGK-IDYDCLIVRFGSSKIDKELINRIRATGQRPHHFLRGI 66
Db 6 TMDDEFTVTPWEY-----KGYDYDCLIVQFGTKTEELKQRIKYLADL-HVTLRRV 59
QY 67 FFSHRDMNOVLDAENKKPFYLYTGRGSSSEAMHVGHLPIFLTKMLQDVFNVPLVIQMT 126
Db 60 FFSHRDLVDLVNDYKSGKFLYTGRAPEL-GHGHGHLPIFLTKMLQKRFNANLYEIT 118
QY 127 DDEKYLK-DLTLDQAGDAVENAKDIIACGFDINKFTIESDLDYMGSSGFYKYNVQ 185
Db 119 DDEKYNRNEFTLIDQRSWAVYDNIIDIVGFPDPTFIFQDEYI---RNMPITVKA 175
QY 186 KHTVFNQVKGIFGFTSDCIKISFPALQAAPSFNSFPQIFRDRDIOCLIPCAIDDP 245
Db 176 KLTTFSEVATRGDLDSSNIGLIFYPALQIAPL-----MPEKK---KCLIPAGIDDP 225
QY 246 YFMTDVAARIGYPRKALHSTFFPALQAQOTKMSADPNSSIFLDTAKQIKTKYKH 305
Db 226 YMLRQDIASIGYTKAAQIHSKFLPPLGPGKSSSNDEPAILVVDPKYERKIMY 285
QY 306 AFGSGRDTIEHRQFGNCVDVVSFMYLTFLEDDDLKQIRKDYTSGAMLTGELKKALI 364
Db 286 AFGSGQPTIELHRKYNPEIDVPFQWLYYFEEEDNRIKEIEEYRSGKMLTGELKQILI 345
QY 365 IEVLQPLIAEHQARRKEVTDEIVKEFMT-PRKLS 397

QY 185 QKVTNQKIGFETDSDICISPAIOAPSFSNPFQIFRDRDIOCLIPCAIDDD 244
 DB 165 -RHATQONVIG-----EPDVGQAFYPAVQTAHL--LPQVHG--EHETLVPIADDD 213
 QY 245 PYRMTDVAIPRIQYP--KEPALHSTFFPALQAGQTKMSADPNSSIFLTDTKQIKTKV 302
 DB 214 PHVRVRDVAAKARYPYKGRGALLMQLPSLAG--PGKMSSS--AGVSIRLTDSPTDYREKV 271
 QY 303 NKIAFSGGRPTIEHROFGNCVDVVSFMVLTFFLEDD--KLEQIKDYTSGAMLTGELK 361
 DB 272 RTAAHYGGRASVEHRAKGVPAEDVPQYLSAFEPEDDAELARIEREYRAGDLSGELK 331
 QY 362 KALIEVLQPIAEHQARKEVTEIYKEFMTPTPLSPD 399
 DB 332 DLAAADRTTEFLAAHQRRRALGD--VTEALDARLITDD 367

RESULT 11

F64476
 tryptophan-tRNA ligase (EC 6.1.1.2) - Methanococcus jannaschii
 N:Alternate names: tryptophanyl-tRNA synthetase
 C:Species: Methanococcus jannaschii
 C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 03-Jun-2002
 C/Accession: F64476
 R:Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirsch, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Tsou, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hultst, M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kahne, B.P.; Borodovsky, M.; Klenk, H.P.; Frazer, C.M.; Smith, H.O.; Woese, C.
 A>Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; MUID:96337999; PMID:868087
 A:Accession: F64476
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-370 <BUL>
 A:Cross-references: GB:U67582; GB:L77117; NID:g1592064; PID:AB99425.1; PID:g1592065; T C:Genetics:
 A:Map position: FOR3375885-1376997
 A:Start codon: GTG
 C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 19.4%; Score 409.5; DB 2; Length 370;
 Best Local Similarity 30.9%; Pred. No. 1,3e-25;
 Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;

QY 17 PWTQTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPHHFLRGIFFSHRMNV 76
 DB 8 PW--ETPAV--IDYKTMEOFGVPIYDVLGLKEE-----HHFRRNITILGHRDERI 57
 QY 77 LDAYENKKPFLYLTGRSPSEAMVGHILPFIPTKMQ--DVFNVLVLOMTDDEKYLW 133
 DB 58 VDAIKNKEFAVVSQMPISGR-MHFGKHMVVDLLKFKYKTDNINIP-----ADLEAYMA 112
 QY 134 KDLTLDOAYGDAV-ENAKDIIACGFDINKTFISDLDYMGSSGFYNNVKI-QKVTFN 191
 DB 113 RNSSFETTKLALINETYITNIALGLDEKINVIQSKYQV-----KDLALLSKTNNWS 167
 QY 192 QVKGIFGTSDSICGISPAIOAPSFSNPFQIFRDRDIOCLIPCAIDDDPYRMT 249
 DB 168 EMKATVGFKEGTNIGHFAIVQVADIL--HPQLDNLSPRPVAVVPGVIDDPIHRL 224
 QY 250 TRVAVR--IGYKPAALLHSTFFPALQAGQTKMSADPNSSIFLTDTKQIKTKV 306
 DB 225 TRDIANNAKEKFTIPESSTYHRFTWGLTG--KMSSSKPEPTALFLTDDEKTKKIKFS-A 281
 QY 307 FSGGRPTIEHROFG--NCDVVSFMVLTFFLEDDKLEQIKDYTSGAMLTGELK 363
 DB 282 KTGKRETLIEHKKYGVPECVVYELFLY--HLILDKEALIEYKGRSSGELTGCKKKA 339
 QY 364 LIEVLOPLIAEHQARKEVTEIYK 388

DB 340 AYERVVEFLKDLKEKREQAETIAVK 364

RESULT 12

E69131
 tryptophan-tRNA ligase (EC 6.1.1.2) - Methanobacterium thermoautotrophicum (strain Delta N:Alternate names: tryptophanyl-tRNA synthetase
 C:Species: Methanobacterium thermoautotrophicum
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 03-Jun-2002
 C/Accession: E69131
 R:Smith, D.R.; Doucette-Stamm, L.A.; DeLonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Iki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: E69131
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-364 <MTH>
 A:Cross-references: GB:AE000812; GB:AE000666; NID:g2621298; PID:AB84757.1; PID:g2621301 C:Genetics:
 A:Gene: MTH251
 A:Start codon: TTG
 C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 18.8%; Score 397.5; DB 2; Length 364;
 Best Local Similarity 27.9%; Pred. No. 1.2e-24;
 Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;

QY 15 VDPWTQTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPHHFLRGIFFSHRMNV 74
 DB 2 IDPW----GSAR-LEIYQDILINFGVRPP-SEVLDVY-----PEPSWLMRGILFGHRDY 50
 QY 75 QVLDAYENKKPFLYLTGRSPSEAMVGHILPFIPTKMQDVFNVLVLOMTDDEKYLW 134
 DB 51 RIISAKKGEDEFAVVTGMPISGR-MHIGKMIYDQLRW-YDRKGAIFIPIDMEAYSR 108
 QY 135 DLTLDOAYGDAV-ENAKDIIACGFDINK-----TFISDLDYMGSSGFYNNVK 183
 DB 109 GVPEDSRRIAIEEYTAGYIALGLDEKDNHIVYLOSENLMWEDLAVV----- 156
 QY 184 IOGVTFENOVKIGFETDSDICISPAIOAPSFSNPFQIFRDRDIOCLIPCAIDDD 243
 DB 157 LACKVFNELRALYGTGTSMAHMTAPILIOVSDILHPQLDELGR---PVIVPGPQD 213
 QY 244 DPEFMTDVAIPRI--GYKPAALLHSTFFPALQAGQTKMSADPNSSIFLTDTKQIK 299
 DB 214 DPHIRLTRDIAARFRDRYGFILPSSSTYHRFMGSLTG--KMSSNRKSAIFLSDTPEEAE 271
 QY 300 TKNKIAFSGGRPTIEHROFGNCVDVVSFMVLTFFLE-DDDKLEQIKDYTSGAMLTG 358
 DB 272 AKI-RNAKTGGRITLKEORELGQVPEECIIYETLLHYMGSSGRLEIYESCRNLTGCG 330
 QY 359 ELKALIEVLQPIAEHQARKE 381
 DB 331 ECKNNTAEIRKFFFEELSVAREK 353

RESULT 13

T43806
 tryptophan-tRNA ligase (EC 6.1.1.2) [imported] - Encephalitozoon cuniculi (fragment)
 C:Species: Encephalitozoon cuniculi
 C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 03-Jun-2002
 C/Accession: T43806
 R:Peyret, A.; Brousseau, V.; Peyret, P.; Metenier, G.; Gouy, M.; Vivares, C.P.
 Mol. Biol. Evol. 15, 683-689, 1998
 A>Title: Microsporidia, antiochondrial protists, possess a 70-kDa heat shock protein gene
 A:Reference number: Z22693; MUID:98277683; PMID:9615449
 A:Accession: T43806
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-134 <PEY>
 A:Cross-references: EMBL:AJ012470; PIDN:CAA10034.1
 C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
 C:Keywords: ligase

Query Match 18.2%; Score 386; DB 2; Length 134;
 Best Local Similarity 51.5%; Pred. No. 2, 5e-24;
 Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

QY 12 EDFVDMVTQTSAAK---GIDYDKLIVRGSSKIDKELINRIERATGQRPHHFLRGIF 67
 DB 3 EGRITFMDVEAVSTDEVPALIDYDKLIINQFGCEKFNQALADRLKSLGKPAHFFRGIV 62
 QY 68 ESHRDNNQVLDAVENKKPPYLTYTGRGSSSEAMHVGHLIFITKMLQDVFNPLVYIQMD 127
 DB 63 FMRDNNLLDELANNRPYLTYTGRGSSKTMHIGHTIFLLCKYQMDFKRLVLIQITD 122
 QY 128 DEKYLKMDLTLD 139
 DB 123 DEKFLKMSWLE 134

RESULT 14

E69461
 tryptophanyl-tRNA synthetase (trps) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
 C:Accession: E69461
 R:Klenk, H.P.; Claydon, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 .; Giodok, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 . Nature 390, 364-370, 1997
 A:Authors: Overbeek, R.; Cotton, M.D.; Spriggs, T.; Artiaeh, P.; Kaite, B.P.; Sykes, S.
 . Smith, H.O.; Woose, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: E69461
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-420 <KLE>
 A:Cross-references: GB:AE000986; GB:AE000782; NID:92689309; PIDN:AAB8954.1; PID:9264885
 C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 17.5%; Score 370.5; DB 2; Length 420;
 Best Local Similarity 27.3%; Pred. No. 2, 3e-22;
 Matches 118; Conservative 60; Mismatches 166; Indels 89; Gaps 12;

QY 15 VDPMTVQTSAAKIDYDKLIVRGSSKIDKELINRIERATGQRPHHFLRGIFFSHRDN 74
 DB 3 VTPWEVEGV---IDYSKLIIEFGMP-SEVLPEID---NPHILMRGALFGHRDY 52
 QY 75 QVLDAVENKKPPYLTYTGRGSSSEAMHVGHLIFITKMLQDVFNPLVYIQMDDEKYLK 134
 DB 53 RILEAKQKPEPVAVMGFMPSG-LPHFGHMTMDELIVMQSAGKAFV-AIDMEHNSR 110
 QY 135 DLTLDQAVGDAVENADIIACGFDINKTIFSDLDYMGSSGFYKVV-KIQKHVTENQY 193
 DB 111 GSWETRELMGLYISIIALGREDVAIVYFOS-----KSHVKKDLAFELSAVNPSEL 164
 QY 194 KGIFGFTDSDCTGKISFPALIQAPSFNSPQIFRDRDIIQCLIPCAIDODPYFRMTRDV 253
 DB 165 RALYGNSTSLAKMVTIQAADIL--HPQLSDGGRKPVVYVGADQDPHMLRLRD 221
 QY 254 APRI----- 257
 DB 222 AARISIFSEPEVGEVVRSRKGAETLSLRLEFDKTIYEEHMDIFGEABETIERAVRKI 281
 QY 258 -----GYPKPALHSTFPALQAGQOTKMSADPNSSIFLDTAQIQTKNVKAFFSG 310
 DB 282 EWEIGGFAPIPPSTYHRTTGLTG- -KWSSSKPEPSYISLDDPREGAKVWK-AFTGG 338
 QY 311 RDTIEHRQFGNCNDVGSFMYLTFLEDD-DKLEQIRKDYTGAMLTGELKKALIEVLQ 369

DB 339 RATAEGRRLGCEPRRCVFEIYSFHLIDSDDELNQLIECEKEGNLLCKCKKMAALVK 398
 QY 370 FLIAEHQARRKEV 382
 DB 399 SFLKEHOKMEAV 411

RESULT 15

D72477

probable tryptophanyl-tRNA synthetase ABE2461 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix
 C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: D72477
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Kaikawa, Y.; Jin-no, K.; Takah
 awa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ka
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: D72477
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-374 <KAW>
 A:Cross-references: DDBJ:AP000064; NID:95105945; PIDN:BAA81476.1; PID:95106165
 A:Experimental source: strain K1
 C:Genetics:
 A:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 16.7%; Score 353; DB 2; Length 374;
 Best Local Similarity 30.4%; Pred. No. 5, 2e-21;
 Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

QY 15 VDPMTVQTSAAKIDYDKLIVRGSSKIDKELINRIERATGQRPHHFLRGIFFSHRDN 74
 DB 8 LDPW-----GAVRIKQYDRILRTFGIRPSEVL--PLRKAGMPEFLMRGIIFGHRDP 61
 QY 75 QVLDAVENKKPPYLTYTGRGSSSEAMHVGHLIFITKMLQ-DVFNPLVYIQMDDEKYLK 133
 DB 62 KLEKKAAGERVAVLTFGPSK-FHFGKLTVDOLIVQKNGFV-FVALDAEAFAV 118
 QY 134 KDLTLDQAVGDAVEN-AKDIACGFDINKT-FISDLDMGSSGFYKVV-KIQKHVTEN 191
 DB 119 RRIIGEEAVRLAEEYIANMILGIDPKDTEYFG-----TNRGTYYFRILQFSGKVTNA 174
 QY 192 QVKGIFG-FTDSDCTGKISFPALIQAPSFNSPQIFRDRDIIQCLIPCAIDODPYFRMT 250
 DB 175 EWEALYIGELTPAKMMASILT---QAADILHVQLDVEGYR--HVVVPVQADQDPHLRLT 227
 QY 251 RVNAPR-----IGYRPALHSTFPALQAGQOTKMSADPNSSIFLDTAKQIKTVNKA 306
 DB 228 RDLADRMAGVLEERPASTYHNLQGLDG--RKMSSSRPDSTIFLDPREVARNKLFR-A 284
 QY 307 PSQGRDITIEHRQFGNCNDV-DVSFMYLTFLEDDKLEQIRKDYTS---GAMLTGELKK 362
 DB 285 LTGGRATAEGRRLGCEPRRCVFEIYSFHLIDSDDELNQLIECEKEGNLLCKCKKMAALVK 398
 QY 363 ALIEVLQPLIAEHQARRKEVTDIVKEFMTPR 394
 DB 342 IAMEKLERFLAEBHQRLKAKTIAMKLVPPR 373

Search completed: July 10, 2003, 12:32:18
 Job time : 16.5539 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:30:15 ; Search time 20.4145 Seconds

(without alignments)
2286.817 Million cell updates/sec

Title: US-09-813-718-10_COPY_71_471

Perfect score: 2116
Sequence: 1 SNHGDPATEAEEDFVDPMTV.....VTDELVKEFMPRKISFPDQ 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2116	100.0	415	9	US-09-813-718-14 Sequence 14, Appl
2	2116	100.0	437	9	US-09-813-718-12 Sequence 12, Appl
3	2116	100.0	471	9	US-10-126-467B-2 Sequence 2, Appl
4	2116	100.0	484	9	US-09-813-718-10 Sequence 10, Appl
5	2101	99.3	471	9	US-09-919-039-163 Sequence 163, Appl
6	2101	99.3	475	10	US-09-925-302-558 Sequence 558, Appl
7	1988	94.0	392	9	US-09-813-718-16 Sequence 16, Appl
8	1125	53.2	433	9	US-10-128-714-8545 Sequence 8545, Ap
9	831	39.3	173	10	US-09-925-302-855 Sequence 855, Ap
10	414.5	19.6	179	9	US-10-128-714-3545 Sequence 3545, Ap
11	328	15.5	85	9	US-09-813-718-45 Sequence 45, Appl
12	292	13.8	85	9	US-09-813-718-46 Sequence 46, Appl
13	273.5	12.9	85	9	US-09-813-718-48 Sequence 48, Appl
14	263	12.4	85	9	US-09-813-718-47 Sequence 47, Appl
15	226.5	10.7	142	10	US-09-925-302-557 Sequence 557, Ap
16	185.5	8.8	341	10	US-09-815-242-13444 Sequence 13444, A
17	147	6.9	46	9	US-09-813-718-51 Sequence 51, Appl
18	147	6.9	385	9	US-10-128-714-3379 Sequence 3379, Ap
19	146.5	6.9	391	9	US-10-128-714-8379 Sequence 8379, Ap

20	146	6.9	339	10	US-09-815-242-11422 Sequence 11422, A
21	142.5	6.7	337	9	US-10-156-761-10954 Sequence 10954, A
22	134	6.3	372	9	US-09-813-718-8 Sequence 8, Appl
23	132.5	6.3	372	9	US-09-813-718-4 Sequence 4, Appl
24	132.5	6.3	334	10	US-09-815-242-11070 Sequence 11070, A
25	128	6.0	334	10	US-09-815-242-11070 Sequence 52, Appl
26	120	5.7	46	9	US-09-813-718-52 Sequence 10678, A
27	113.5	5.4	423	10	US-09-815-242-10678 Sequence 5494, Ap
28	113	5.3	415	10	US-09-815-242-5494 Sequence 12649, A
29	113	5.3	420	10	US-09-815-242-12649 Sequence 4252, Ap
30	112	5.3	345	9	US-09-738-626-4252 Sequence 12259, A
31	110.5	5.2	331	9	US-10-156-761-12259 Sequence 50, Appl
32	108	5.1	306	9	US-10-126-927-50 Sequence 59, Appl
33	108	5.1	306	9	US-10-126-927-59 Sequence 64, Appl
34	108	5.1	306	9	US-10-126-927-64 Sequence 50, Appl
35	108	5.1	306	9	US-10-126-931A-50 Sequence 59, Appl
36	108	5.1	306	9	US-10-126-931A-59 Sequence 64, Appl
37	108	5.1	306	9	US-10-126-931A-64 Sequence 13806, A
38	108	5.1	334	10	US-09-815-242-13806 Sequence 51, Appl
39	107	5.1	306	9	US-10-126-927-51 Sequence 51, Appl
40	107	5.1	306	9	US-10-126-931A-51 Sequence 10371, A
41	106.5	5.0	334	10	US-09-815-242-10371 Sequence 45, Appl
42	106	5.0	306	9	US-10-126-927-45 Sequence 54, Appl
43	106	5.0	306	9	US-10-126-931A-45 Sequence 54, Appl
44	105.5	5.0	46	9	US-09-813-718-54 Sequence 43, Appl
45	105	5.0	306	9	US-10-126-927-43 Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-09-813-718-14
Sequence 14, Application US/09813718
Publication No. US2002018266A1
GENERAL INFORMATION:
APPLICANT: Schimmel, Paul
APPLICANT: Makasugi, Keiuke
TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
FILE REFERENCE: 00-221
CURRENT APPLICATION NUMBER: US/09/813,718
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 415
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: human
OTHER INFORMATION: supermatl tips in pET20B
US-09-813-718-14

Query Match 100.0%; Score 2116; DB 9; Length 415;
Best Local Similarity 100.0%; Pred. No. 5.7e-192;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	SNHGDPATEAEEDFVDPMTVQTSSAKGIDYDKLIVRFSSSKIDKELINRIEATQGRPHH	60
DB	2	SNHGDPATEAEEDFVDPMTVQTSSAKGIDYDKLIVRFSSSKIDKELINRIEATQGRPHH	61
QY	61	FLRRGIFPSHRMNOVLAYENKKPFYLYTGSPSEAMHVGHLPIFTKMLQVFNVP	120
DB	62	FLRRGIFPSHRMNOVLAYENKKPFYLYTGSPSEAMHVGHLPIFTKMLQVFNVP	121
QY	121	LVIQMTDEKYLKMLTDQAYGAVENAKDIIACGPDINKTIFPSDIDYMGSSGFYKN	180
DB	122	LVIQMTDEKYLKMLTDQAYGAVENAKDIIACGPDINKTIFPSDIDYMGSSGFYKN	181
QY	181	VVKIQHVTFNQKGI FGFTSDCI GIKISFPAIQAPSPNSFPQIFRDRTDIQCLIPCA	240
DB	182	VVKIQHVTFNQKGI FGFTSDCI GIKISFPAIQAPSPNSFPQIFRDRTDIQCLIPCA	241

```

QY      241 IDDDPYFRMTDVAPRIGYKPPALHSTFPFPAIOGAQTMSASDPNSSIFLTTAKOIKT 300
      242 IDDDPYFRMTDVAPRIGYKPPALHSTFPFPAIOGAQTMSASDPNSSIFLTTAKOIKT 301
QY      301 KVNKAHAFSGGRDTEIEHRQFGNCDDVVSFMYLTFPLEDDDKLEQIRKDYTSGAMLTGEL 360
      302 KVNKAHAFSGGRDTEIEHRQFGNCDDVVSFMYLTFPLEDDDKLEQIRKDYTSGAMLTGEL 361
QY      361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401
      362 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 402

```

```

RESULT 2
US-09-813-718-12
; Sequence 12, Application US/09813718
; Publication No. US20020182666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakasugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; TITLE OF INVENTION: The Regulation of Angiogenesis
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813, 718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 437
; TYPE: PRF
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: human mini
; OTHER INFORMATION: Trps in per208
US-09-813-718-12

```

```

Query Match      100.0%; Score 2116; DB 9; Length 437;
Best Local Similarity 100.0%; Pred. No. 6.1e-192; Indels 0; Gaps 0;
Matches 401; Conservative 0; Mismatches 0;

```

```

QY      1 SNHGPDATAEEDFVDPWTVQTSASAKIDYDKLIVRFGSSKIDKELINRIERATGQRPH 60
      24 SNHGPDATAEEDFVDPWTVQTSASAKIDYDKLIVRFGSSKIDKELINRIERATGQRPH 83
QY      61 FLRRGIFFSHRDNNQVLDAVENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVFNVP 120
      84 FLRRGIFFSHRDNNQVLDAVENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVFNVP 143
QY      121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180
      144 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 203
QY      181 VVKIQKAVTFNQVKGIFGFTDSDCIKISFPALIOAAPSFSNPFQIFRDRDIOCLIPCA 240
      204 VVKIQKAVTFNQVKGIFGFTDSDCIKISFPALIOAAPSFSNPFQIFRDRDIOCLIPCA 263
QY      241 IDDDPYFRMTDVAPRIGYKPPALHSTFPFPAIOGAQTMSASDPNSSIFLTTAKOIKT 300
      264 IDDDPYFRMTDVAPRIGYKPPALHSTFPFPAIOGAQTMSASDPNSSIFLTTAKOIKT 323
QY      301 KVNKAHAFSGGRDTEIEHRQFGNCDDVVSFMYLTFPLEDDDKLEQIRKDYTSGAMLTGEL 360
      324 KVNKAHAFSGGRDTEIEHRQFGNCDDVVSFMYLTFPLEDDDKLEQIRKDYTSGAMLTGEL 383
QY      361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401
      384 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 424

```

```

RESULT 3
US-10-126-467B-2
; Sequence 2, Application US/10126467B

```

```

; Publication NO. US20030059797A1
; GENERAL INFORMATION:
; APPLICANT: Paley, Elena
; TITLE OF INVENTION: ANIMAL MODEL OF AND TEST FOR ALZHEIMER'S DISEASE
; FILE REFERENCE: PAL-111
; CURRENT APPLICATION NUMBER: US/10/126,467B
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/284,980
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 09/513,895
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 09/384,869
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 471
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-126-467B-2

```

```

Query Match      100.0%; Score 2116; DB 9; Length 471;
Best Local Similarity 100.0%; Pred. No. 6.8e-192; Indels 0; Gaps 0;
Matches 401; Conservative 0; Mismatches 0;

```

```

QY      1 SNHGPDATAEEDFVDPWTVQTSASAKIDYDKLIVRFGSSKIDKELINRIERATGQRPH 60
      71 SNHGPDATAEEDFVDPWTVQTSASAKIDYDKLIVRFGSSKIDKELINRIERATGQRPH 130
QY      61 FLRRGIFFSHRDNNQVLDAVENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVFNVP 120
      131 FLRRGIFFSHRDNNQVLDAVENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVFNVP 190
QY      121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180
      191 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 250
QY      181 VVKIQKAVTFNQVKGIFGFTDSDCIKISFPALIOAAPSFSNPFQIFRDRDIOCLIPCA 240
      251 VVKIQKAVTFNQVKGIFGFTDSDCIKISFPALIOAAPSFSNPFQIFRDRDIOCLIPCA 310
QY      241 IDDDPYFRMTDVAPRIGYKPPALHSTFPFPAIOGAQTMSASDPNSSIFLTTAKOIKT 300
      311 IDDDPYFRMTDVAPRIGYKPPALHSTFPFPAIOGAQTMSASDPNSSIFLTTAKOIKT 370
QY      301 KVNKAHAFSGGRDTEIEHRQFGNCDDVVSFMYLTFPLEDDDKLEQIRKDYTSGAMLTGEL 360
      371 KVNKAHAFSGGRDTEIEHRQFGNCDDVVSFMYLTFPLEDDDKLEQIRKDYTSGAMLTGEL 430
QY      361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401
      431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 471

```

```

RESULT 4
US-09-813-718-10
; Sequence 10, Application US/09813718
; Publication No. US20020182666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakasugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; TITLE OF INVENTION: The Regulation of Angiogenesis
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813, 718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 484
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:

```

OTHER INFORMATION: Description of Artificial Sequence: human
OTHER INFORMATION: full-length Trps in pET20B
US-09-813-718-10

Query Match 100.0%; Score 2116; DB 9; Length 484;
Best Local Similarity 100.0%; Pred. No. 7.1e-192;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGDATEAEEDFVDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRERATGQRPH 60
DB 71 SNHGDATEAEEDFVDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRERATGQRPH 130
QY 61 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 120
DB 131 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 190
QY 121 LVIQMTDBEKYLMKDLTDQAYGDAVENAKDIIACGFDINKTPIFSDDLDMGMSGGFYKN 180
DB 191 LVIQMTDBEKYLMKDLTDQAYGDAVENAKDIIACGFDINKTPIFSDDLDMGMSGGFYKN 250
QY 181 VKIKQHTVFNQVKIGFPTSDCIGKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 240
DB 251 VKIKQHTVFNQVKIGFPTSDCIGKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 310
QY 241 IDDDYFEMTRDVAPRIGYKPKALHSTFPALQAGQTKMSASDPNSSIFLDTAKQIKT 300
DB 311 IDDDYFEMTRDVAPRIGYKPKALHSTFPALQAGQTKMSASDPNSSIFLDTAKQIKT 370
QY 301 KVNKAIFSGGRDTIEHRQFGNCDVDVSFWYLFTELEDDDKLEQIRKDYTSGAMLTGEL 360
DB 371 KVNKAIFSGGRDTIEHRQFGNCDVDVSFWYLFTELEDDDKLEQIRKDYTSGAMLTGEL 430
QY 361 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401
DB 431 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 471

RESULT 5

US-09-919-039-163
Sequence 163, Application US/09919039
Publication No. US20030108871A1
GENERAL INFORMATION:
APPLICANT: Kaser, Matthew R.
TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
FILE REFERENCE: PA-0035 US
CURRENT FILING DATE: 2002-09-09
PRIOR APPLICATION NUMBER: 60/222,113
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 401
SOFTWARE: PERL Program
SEQ ID NO 163
LENGTH: 471
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030108871A1 2705515CD1
US-09-919-039-163

Query Match 99.3%; Score 2101; DB 9; Length 471;
Best Local Similarity 99.5%; Pred. No. 1.8e-190;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNHGDATEAEEDFVDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRERATGQRPH 60
DB 71 SNHGDATEAEEDFVDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRERATGQRPH 130
QY 61 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 120
DB 131 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 190
QY 121 LVIQMTDBEKYLMKDLTDQAYGDAVENAKDIIACGFDINKTPIFSDDLDMGMSGGFYKN 180

DB 191 LVIQMTDBEKYLMKDLTDQAYGDAVENAKDIIACGFDINKTPIFSDDLDMGMSGGFYKN 250
QY 181 VKIKQHTVFNQVKIGFPTSDCIGKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 240
DB 251 VKIKQHTVFNQVKIGFPTSDCIGKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 310
QY 241 IDDDYFEMTRDVAPRIGYKPKALHSTFPALQAGQTKMSASDPNSSIFLDTAKQIKT 300
DB 311 IDDDYFEMTRDVAPRIGYKPKALHSTFPALQAGQTKMSASDPNSSIFLDTAKQIKT 370
QY 301 KVNKAIFSGGRDTIEHRQFGNCDVDVSFWYLFTELEDDDKLEQIRKDYTSGAMLTGEL 360
DB 371 KVNKAIFSGGRDTIEHRQFGNCDVDVSFWYLFTELEDDDKLEQIRKDYTSGAMLTGEL 430
QY 361 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401
DB 431 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 471

RESULT 6

US-09-925-302-558
Sequence 558, Application US/09925302
Patent No. US20020044941A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA104
CURRENT APPLICATION NUMBER: US/09/925,302
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05918
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 896
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 558
LENGTH: 475
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-302-558

Query Match 99.3%; Score 2101; DB 10; Length 475;
Best Local Similarity 99.5%; Pred. No. 1.8e-190;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNHGDATEAEEDFVDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRERATGQRPH 60
DB 75 SNHGDATEAEEDFVDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRERATGQRPH 134
QY 61 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 120
DB 135 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 194
QY 121 LVIQMTDBEKYLMKDLTDQAYGDAVENAKDIIACGFDINKTPIFSDDLDMGMSGGFYKN 180
DB 195 LVIQMTDBEKYLMKDLTDQAYGDAVENAKDIIACGFDINKTPIFSDDLDMGMSGGFYKN 254
QY 181 VKIKQHTVFNQVKIGFPTSDCIGKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 240
DB 255 VKIKQHTVFNQVKIGFPTSDCIGKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 314
QY 241 IDDDYFEMTRDVAPRIGYKPKALHSTFPALQAGQTKMSASDPNSSIFLDTAKQIKT 300
DB 315 IDDDYFEMTRDVAPRIGYKPKALHSTFPALQAGQTKMSASDPNSSIFLDTAKQIKT 374
QY 301 KVNKAIFSGGRDTIEHRQFGNCDVDVSFWYLFTELEDDDKLEQIRKDYTSGAMLTGEL 360
DB 375 KVNKAIFSGGRDTIEHRQFGNCDVDVSFWYLFTELEDDDKLEQIRKDYTSGAMLTGEL 434
QY 361 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401
DB 435 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 475

```

RESULT 7
US-09-813-718-16
; Sequence 16; Application US/09813718
; Publication No. US20020182666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakauegi, Keisuke
; TITLE OF INVENTION: Human Aminocycl-terNA Synthetase Polypeptides Useful For
; TITLE OF INVENTION: The Regulation of Angiogenesis
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentin Ver. 2.0
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: human minor
; OTHER INFORMATION: Trps fragment in pET20B
US-09-813-718-16

Query Match          94.0%; Score 1988; DB 9; Length 392;
Best Local Similarity 100.0%; Pred. No. 6,9e-180;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 SAKGIDYDKLIVRFGSSKIDKELINRTERATGCRPHHFRGIFFSHRDNNQVLDAYENK 83
DB 2 SAKGIDYDKLIVRFGSSKIDKELINRTERATGCRPHHFRGIFFSHRDNNQVLDAYENK 61
QY 84 KPEYLYTGRGSSSEAMHVGHLIPFTKWLQDVFNVLVITQMTDEKYLTKDLTDQAYG 143
DB 62 KPEYLYTGRGSSSEAMHVGHLIPFTKWLQDVFNVLVITQMTDEKYLTKDLTDQAYG 121
QY 144 DAVENKADIIACGPDINKTFIFSDLDYMGSSGKYNVVKIQGVFNQVKGIFGFDSD 203
DB 122 DAVENKADIIACGPDINKTFIFSDLDYMGSSGKYNVVKIQGVFNQVKGIFGFDSD 181
QY 204 CIKISFPALQAPSPFSNSFPQIFRDRTDIQCILPCAIDDPYFRMTRDVAAPRIGYKPA 263
DB 182 CIKISFPALQAPSPFSNSFPQIFRDRTDIQCILPCAIDDPYFRMTRDVAAPRIGYKPA 241
QY 264 LHMSTFPALQAGQTKMSASDPNSSIFLDTAKQIKTKVKNHAFSGGRDTIEHRQFGCN 301
DB 242 LHMSTFPALQAGQTKMSASDPNSSIFLDTAKQIKTKVKNHAFSGGRDTIEHRQFGCN 301
QY 324 CDVDVSPFMYLTFFLEDDDKLEQIRKQVTSGLMTGELKKALIEVLOPLIAEHQARRKEVT 383
DB 302 CDVDVSPFMYLTFFLEDDDKLEQIRKQVTSGLMTGELKKALIEVLOPLIAEHQARRKEVT 361
QY 384 DEIVKEFMTPRKLSFDFQ 401
DB 362 DEIVKEFMTPRKLSFDFQ 379

RESULT 8
US-10-128-714-8545
; Sequence 8545; Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroszhin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714

```

```

; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 8545
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8545

Query Match          53.2%; Score 1125; DB 9; Length 433;
Best Local Similarity 54.7%; Pred. No. 4,2e-98;
Matches 220; Conservative 63; Mismatches 107; Indels 12; Gaps 5;

QY 7 ATEAEEDFVDPMTV-----QTSASAKGIDYDKLIVRFGSSKIDKELINRTERATGCRPHH 61
DB 22 ASKAVAQVYTPFDVSGVDSGSKLPVDYDKLVREGRGARRISKELLERERVTGRPHH 81
QY 62 LRRGIFFSHRDNNQVLDAYENKPPYLYTGRGSSSEAMHVGHLIPFTKWLQDVFNVLV 121
DB 82 MRRGIFFSHRDNLILIDRYEKGQPFYLYTGRGSSSEAMHVGHLIPFTKWLQDVFNVLV 141
QY 122 VIQMTDEKYLTKDLTDQAYENKADIIACGPDINKTFIFSDLDYMGSSGKYNVVK 180
DB 142 VIQMTDEKYLTKDLTDQAYENKADIIACGPDINKTFIFSDLDYMGSSGKYNVVK 199
QY 181 VVKIQGVFNQVKGIFGFDSDCIKISFPALQAPSPFSNSFPQIF-RDR--TDIOCL 236
DB 200 ICMARIRIINSVRGFGFNDNNVGEHFCAQTQATATSPHIFGDRKKVSSIPCL 259
QY 237 IPCAIDDPYFRMTRDVAAPRIGYKPAIHSFPALQAGQTKMSASDPNSSIFLDTAK 296
DB 260 IPCAIDDPYFRMTRDVAAPRIGYKPAIHSFPALQAGQTKMSASDPNSSIFLDTAK 319
QY 297 QIKTKVKNHAFSGGRDTIEHRQFGCNCDVDVSPFMYLTFFLEDDDKLEQIRKQVTSGLMT 356
DB 320 RINKKINKYAFSGGQTAELQROLGANTYKQDVFPQYLTFFMEDDELIRIVAYEKGEM 379
QY 357 TGELEKALIEVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSF 398
DB 380 TGEVKKQICIAELQAVYQAFQERRAQVTDIVAEFMRPRSLFW 421

RESULT 9
US-09-925-302-855
; Sequence 855; Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 855
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```


NAME/KEY: SITE
LOCATION: (159)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (168)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-855

Query Match
Best Local Similarity 39.3%; Score 831; DB 10; Length 173;
Pred. No. 7.9e-71;
Matches 156; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 107 FIFTKMLQOVENVPLVQMTDEKYLMDLTLDQAYGDAVENAKDIACGFDINKTFIRS 166
DB 1 FIFTKMLQOVENVPLVQMTDEKYLMDLTLDQAYGDAVENAKDIACGFDINKTFIRS 60
QY 167 DLDYMGSSGFYKNNVKIQKHTFNOVKIGFTSDCIGKISFPAIQAPSPNSFPQI 226
DB 61 DLDYMGSSGFYKNNVKIQKHTFNOVKIGFTSDCIGKISFPAIQAPSPNSFPQI 120
QY 227 FRDRTDIQCILPCALIDODPEFRMTKRVAPRIGYKPKAL 264
DB 121 FRDRTDIQCILPCALIDODPEFRMTKRVAPRIGYKPKAL 158

RESULT 10
US-10-128-714-3545

Sequence 3545, Application US/10128714
Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Weng
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Broshkin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128, 714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285, 697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287, 066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295, 890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303, 899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316, 362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3545
LENGTH: 179
TYPE: PRT
ORGANISM: Aspergillus fumigatus
US-10-128-714-3545

Query Match
Best Local Similarity 19.6%; Score 414.5; DB 9; Length 179;
Pred. No. 2.3e-11;
Matches 76; Conservative 20; Mismatches 27; Indels 5; Gaps 1;

QY 7 ATBAEDFVDPMTV-----QTSKAGIDYDKLIVRGSSKIDKELINRIERATGQRPHP 61
DB 22 ASKAVAVQVTPPVDVSGVDSGLLPVDKLVREBGAIRISKELLERFERVTGRRPFR 81
QY 62 LRRGIFFSHRDMNQVDAVENKKPFYLYTGRGSSSEAMVGHILPFIPTKMLQOVENVPL 121
DB 82 MRGIVFSHRDLNLIIDRIEKGOPFYLYTGRGSSSEAMVGHITPPEFTKMLQOVENVPL 141
QY 122 VIQMTDDE 129
||:|||||

DB 142 VIQMTDDE 149

RESULT 11
US-09-813-718-45

Sequence 45, Application US/09813718
Publication No. US2002018266A1
GENERAL INFORMATION:
APPLICANT: Schimmel, Paul
APPLICANT: Wakasugi, Keisuke
TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
FILE REFERENCE: 00-221
CURRENT APPLICATION NUMBER: US/09/813, 718
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 45
LENGTH: 85
TYPE: PRT
ORGANISM: Homo sapiens
US-09-813-718-45

Query Match
Best Local Similarity 15.5%; Score 328; DB 9; Length 85;
Pred. No. 1.2e-23;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATBAEDFVDPMTVQTSKAGIDYDKLIVRGSSKIDKELINRIERATGQRPHP 60
DB 24 SNHGPDATBAEDFVDPMTVQTSKAGIDYDKLIVRGSSKIDKELINRIERATGQRPHP 83

QY 61 FL 62
DB 84 FL 85

RESULT 12
US-09-813-718-46

Sequence 46, Application US/09813718
Publication No. US2002018266A1
GENERAL INFORMATION:
APPLICANT: Schimmel, Paul
APPLICANT: Wakasugi, Keisuke
TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
FILE REFERENCE: 00-221
CURRENT APPLICATION NUMBER: US/09/813, 718
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 46
LENGTH: 85
TYPE: PRT
ORGANISM: Bos taurus
US-09-813-718-46

Query Match
Best Local Similarity 13.8%; Score 292; DB 9; Length 85;
Pred. No. 3.2e-20;
Matches 56; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 SNHGPDATBAEDFVDPMTVQTSKAGIDYDKLIVRGSSKIDKELINRIERATGQRPHP 60
DB 24 SGEGLDATABEDFVDPMTVQTSKAGIDYDKLIVRGSSKIDKELINRIERATGQRPHP 83

QY 61 FL 62
DB 84 FL 85

RESULT 13
US-09-813-718-48
Sequence 48, Application US/09813718
Publication No. US2002018266A1

```

; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 48
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-813-718-48

Query Match      12.9%; Score 273.5; DB 9; Length 85;
Best Local Similarity 83.9%; Pred. No. 1.8e-18;
Matches 52; Conservative 6; Mismatches 3; Indels 1; Gaps 1;

QY      2 NHGPD-ATEABEDFVDPVPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 60
Db      24 SHGDEAVDKEDEFDVPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 83
QY      61 FL 62
Db      84 FL 85

RESULT 14
US-09-813-718-47
; Sequence 47, Application US/09813718
; Publication No. US2002018266A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 47
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-813-718-47

Query Match      12.4%; Score 263; DB 9; Length 85;
Best Local Similarity 85.2%; Pred. No. 1.8e-17;
Matches 52; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      2 NHGPD-ATEABEDFVDPVPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 61
Db      25 NCSDATKASEDFVDPVPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 84
QY      62 L 62
Db      85 L 85

RESULT 15
US-09-925-302-557
; Sequence 557, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
```

```

; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 557
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (122)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (117)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (130)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (137)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-557

Query Match      10.7%; Score 226.5; DB 10; Length 142;
Best Local Similarity 70.6%; Pred. No. 1.1e-13;
Matches 48; Conservative 5; Mismatches 12; Indels 3; Gaps 2;

QY      1 SNHGPD-ATEABEDFVDPVPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 60
Db      75 SNHGPD-ATEABEDFVDPVPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 133
QY      61 FLRGIFF 68
Db      134 --GQGXFF 139
```

Search completed: July 10, 2003, 12:33:37
Job time: 21.4145 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model1

Run on: July 10, 2003, 12:29:05 ; Search time 12.1515 Seconds
(without alignments)
970.956 Million cell updates/sec

Title: US-09-813-718-10_COPY_71_471
Perfect score: 2116
Sequence: 1 SNHGPDATAEEDFPDPTV.....VTDEIVKEMTRKLSFDQ 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

1: Issued Patents AA:*
2: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/6C_COMB.pep:*
7: /cgn2_6/prodata/1/iaa/6D_COMB.pep:*
8: /cgn2_6/prodata/1/iaa/6E_COMB.pep:*
9: /cgn2_6/prodata/1/iaa/6F_COMB.pep:*
10: /cgn2_6/prodata/1/iaa/6G_COMB.pep:*
11: /cgn2_6/prodata/1/iaa/6H_COMB.pep:*
12: /cgn2_6/prodata/1/iaa/6I_COMB.pep:*
13: /cgn2_6/prodata/1/iaa/6J_COMB.pep:*
14: /cgn2_6/prodata/1/iaa/6K_COMB.pep:*
15: /cgn2_6/prodata/1/iaa/6L_COMB.pep:*
16: /cgn2_6/prodata/1/iaa/6M_COMB.pep:*
17: /cgn2_6/prodata/1/iaa/6N_COMB.pep:*
18: /cgn2_6/prodata/1/iaa/6O_COMB.pep:*
19: /cgn2_6/prodata/1/iaa/6P_COMB.pep:*
20: /cgn2_6/prodata/1/iaa/6Q_COMB.pep:*
21: /cgn2_6/prodata/1/iaa/6R_COMB.pep:*
22: /cgn2_6/prodata/1/iaa/6S_COMB.pep:*
23: /cgn2_6/prodata/1/iaa/6T_COMB.pep:*
24: /cgn2_6/prodata/1/iaa/6U_COMB.pep:*
25: /cgn2_6/prodata/1/iaa/6V_COMB.pep:*
26: /cgn2_6/prodata/1/iaa/6W_COMB.pep:*
27: /cgn2_6/prodata/1/iaa/6X_COMB.pep:*
28: /cgn2_6/prodata/1/iaa/6Y_COMB.pep:*
29: /cgn2_6/prodata/1/iaa/6Z_COMB.pep:*
30: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
31: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
32: /cgn2_6/prodata/1/iaa/6C_COMB.pep:*
33: /cgn2_6/prodata/1/iaa/6D_COMB.pep:*
34: /cgn2_6/prodata/1/iaa/6E_COMB.pep:*
35: /cgn2_6/prodata/1/iaa/6F_COMB.pep:*
36: /cgn2_6/prodata/1/iaa/6G_COMB.pep:*
37: /cgn2_6/prodata/1/iaa/6H_COMB.pep:*
38: /cgn2_6/prodata/1/iaa/6I_COMB.pep:*
39: /cgn2_6/prodata/1/iaa/6J_COMB.pep:*
40: /cgn2_6/prodata/1/iaa/6K_COMB.pep:*
41: /cgn2_6/prodata/1/iaa/6L_COMB.pep:*
42: /cgn2_6/prodata/1/iaa/6M_COMB.pep:*
43: /cgn2_6/prodata/1/iaa/6N_COMB.pep:*
44: /cgn2_6/prodata/1/iaa/6O_COMB.pep:*
45: /cgn2_6/prodata/1/iaa/6P_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1218.5	57.6	424	4	US-08-876-885-26
2	185.5	8.8	341	3	US-08-928-100-2
3	185.5	8.8	341	4	US-09-492-581-2
4	185.5	8.8	341	4	US-09-425-666-2
5	164.5	7.8	409	2	US-08-743-130A-39
6	162.5	7.7	409	2	US-08-743-130A-2
7	132.5	6.3	388	3	US-08-705-868-4
8	132.5	6.3	388	3	US-09-123-615-4
9	113.5	5.2	418	4	US-08-855-910-11
10	109	5.2	377	4	US-09-352-990-28
11	108	5.1	197	2	US-08-923-867-2
12	108	5.1	197	3	US-08-928-100-4
13	108	5.1	197	4	US-09-183-134-2
14	108	5.1	197	4	US-09-492-581-4
15	108	5.1	197	4	US-09-425-666-4
16	104.5	4.9	370	2	US-08-415-593-45
17	100.5	4.7	427	4	US-09-134-001C-5141
18	95	4.5	418	3	US-08-844-054-2
19	95	4.5	418	4	US-09-347-333-2
20	95	4.5	464	4	US-09-134-001C-4701
21	94.5	4.5	877	2	US-08-907-166-8
22	89.5	4.2	344	4	US-09-393-554-2
23	88.5	4.2	501	4	US-09-157-257-8
24	88.5	4.2	898	1	US-08-465-995A-4
25	88.5	4.2	898	2	US-08-465-994C-4
26	88.5	4.2	898	2	US-08-966-145-4
27	88.5	4.2	920	1	US-08-101-593-4

ALIGNMENTS

28	88	4.2	606	2	US-08-883-534-3	Sequence 3, Appl
29	88	4.2	606	3	US-09-204-764-3	Sequence 3, Appl
30	87.5	4.1	855	4	US-08-890-865A-10	Sequence 10, Appl
31	87.5	4.1	898	1	US-08-465-995A-2	Sequence 2, Appl
32	87.5	4.1	898	2	US-08-465-994C-2	Sequence 2, Appl
33	87.5	4.1	898	2	US-08-966-145-2	Sequence 2, Appl
34	87.5	4.1	920	1	US-08-101-593-2	Sequence 2, Appl
35	87	4.1	502	4	US-09-134-001C-4511	Sequence 4511, Ap
36	86.5	4.1	849	4	US-09-157-257-4	Sequence 4, Appl
37	86	4.1	930	4	US-09-134-001C-5314	Sequence 5314, Ap
38	85	4.0	410	4	US-09-352-990-16	Sequence 16, Appl
39	84.5	4.0	539	3	US-09-157-257-6	Sequence 6, Appl
40	83.5	3.9	428	3	US-08-331-625A-43	Sequence 43, Appl
41	83.5	3.9	428	4	US-09-494-153-43	Sequence 43, Appl
42	83.5	3.9	510	1	US-08-249-112-3	Sequence 3, Appl
43	83.5	3.9	510	5	PCT-US95-06556-3	Sequence 3, Appl
44	83.5	3.9	970	1	US-08-375-709-7	Sequence 7, Appl
45	83.5	3.9	970	1	US-08-752-929-7	Sequence 7, Appl

RESULT 1
US-08-876-885-26
Sequence 26, Application US/08876885
Patent No. 6174713

GENERAL INFORMATION:
APPLICANT: Shen, Xiaoyu
TITLE OF INVENTION: CANDIDA CYTOPLASMIC TRYPTOPANYL-tRNA
TITLE OF INVENTION: SYNTHETASE PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING
TITLE OF INVENTION: SAME
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/876,885
FILING DATE: 16-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI97-02
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
US-08-876-885-26

Query Match 57.6%; Score 1218.5; DB 4; Length 424;
Best Local Similarity 57.1%; Pred. No. 3.7e-128;
Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

QY 8 TAEEDFPDPTVQ---TSSAKGIDYKLVFGSSKIDKELNIRIRATGDPHPFLR 63
DB 13 TEESEKQITPMEVEGAVVDGKSMGIDYDKLISQFGTKHITETTLERFQVGTGEHPFLK 72

QY 64 RGIFSHRDNMVLDVAENKKPFYLTGRRSSSEAMVGHILPFTKMLQDVENVPLVT 123
DB 73 RGFFSRDRIDRIDLYEHGEPFLYTGRRSSSMHGHVPIFTKMLQDVENVPLVT 132
QY 124 QMTDEKYLWK-DLTLDQAYGDAVENAKDIJACGFDINKTIFESDLDYMGMSGFYKNV 182
DB 133 ELTDEKFLPHQTLTIDVKGFAENAKDIJAVGFNENPFIFSDLYMG--GAFYENVV 190
QY 183 KICKHFNQVKGIFGFTSDCTGKISFPALQAPSPNSFPQIFRRTDIOCLIPCAID 242
DB 191 RTRQITTSYAKAVGFTSDCTGKIHFASTIQTAFSPSSPDVGLPPTKCLIPCAID 250
QY 243 QDVFYRTRDVAPRIGYKPKALLHSTFPALQAGTMSADPNSSIFLDTAKQIKTV 302
DB 251 QDVFYRTRDVADKLTFTKALHAKFPALQAGTMSADPNSSIFLMDTAKQIKTV 310
QY 303 NKHAFSGGRDTIEHRQFGNCVDVSPMYLTFLEDDDKLEQIRKDYTSGAMLTGELKK 362
DB 311 NKYAFSGGRATAEHRRELGNPEVDVAFQYLSFYSYDEKLAQLQEGYRKGEILSGMKK 370
QY 363 ALIEVLOPLAEHQARKEVTDIEVKEFMTPRKLSF 398
DB 371 ECITVLOEFVSAQERRSKVDQVEKEMKPKLVF 406

RESULT 2

US-08-928-100-2

Sequence 2, Application US/08928100

Patent No. 6046174

GENERAL INFORMATION:

APPLICANT: Gentry, Danile

APPLICANT: Greenwood, Claire

TITLE OF INVENTION: No. 6046174el trps

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/928,100

FILING DATE: 12-SEP-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 9619072.3

FILING DATE: 12-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Gimmil, Edward R

REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P31624-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 341 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-928-100-2

Query Match

Best Local Similarity

8.8%; Score 185.5; DB 3; Length 341;

24.0%; Pred. No. 3.2e-12;

Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;
QY 84 KPPYLYTGRRSSSEAMVGHILPFTKMLQDVENVPLVTQMTDEK-LMKDLTLDQAY 142
DB 3 KPYIL-TGDRPTOK-HIHYVGSILNR-----VLQEBKYMFPFLADQAL 49
QY 143 GDVENV-----AKDIJACGFDINKTIF--SDLDYMGMSGFYKNV---KICK 186
DB 50 TDHAKPQITVBSIGNVALDYLAAGDPKSTIFITOSQIPELAEISMTYVLSLARLR 109
QY 187 HTFNQVKGIFGFTSDCTGKISFPALQAA--PSFNSFPQIFRRTDIOCLIPCAID 244
DB 110 NPTVKTEISQKGEISIPGFLVYPIAQAADITAFKAY-----VPVGTQK 156
QY 245 PYRMTRD-----VAPRIGYKPKALLHSTFPALQAGTMSADPNSSIF 290
DB 157 PMIQREIVRSFNANVCDVLPVPEGIPENE--RAGLPGLDG--NAKMSKS--LNGIY 212
QY 291 LPTDAKQIKTVKNKHAFAFGGRDTIEHRQFGNCVDVSPMYLTFE--LEDDDKLEQIR 348
DB 213 LADDAITLKKVMSMTTDPDHIREVDPKIEGN---WVFHLDVGRFEDPAQELADME 268
QY 349 DYTSGAMLTGELKALIEVLOPLAEHQARKEVTDIE 386
DB 269 RYORGGLGVKTRVYLITELERELGPIRRRIEFAKDM 306

RESULT 3

US-09-492-581-2

Sequence 2, Application US/09492581

Patent No. 6346409

GENERAL INFORMATION:

APPLICANT: Gentry, Danile

APPLICANT: Greenwood, Claire

TITLE OF INVENTION: No. 6346409el trps

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/492,581

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/928,100

FILING DATE: 12-SEP-1997

APPLICATION NUMBER: 9619072.3

ATTORNEY/AGENT INFORMATION:

NAME: Gimmil, Edward R

REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P31624-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 341 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-492-581-2

Query Match 8.8%; Score 185.5; DB 4; Length 341;
Best Local Similarity 24.0%; Pred. No. 3.2e-12;
Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 84 KPEVLTGKSPSEAMHGHLPFTKMLQDVFNVLQMTDDEKY-LMKDLTLDDQAY 142
DB 3 KPILL-TGDRPTCK-LHIGHVGSLSKNR-----VLLQEBKXDMFVFLADQQL 49
QY 143 GDAVEN-----AKDIIACGFDINKTFIF--SDLDYMGSSGFYKXNV---KIOK 186
DB 50 THAKDPQITVESIGNVALDYLAAGLDPKSTIFIQSQIPELAEISMTYMLVSLARLR 109
QY 187 HVTENQVKIGFPTSDCIKISFPALQAA--PSFNSFPQIFRDRTDIQCILPCALDD 244
DB 110 NPTVKTETISOKGFGESIPFGFLVYPLAQADITAFKANY-----VPVGTQK 156
QY 245 PYFRMTRD-----VAPRIGYKPKALHSTFPALQGAQTMSASDPNSIF 290
DB 157 PMIEQIREIVRSFNNAVNCVIVPEGIYPENE--RAGRLPGIDG--NAKMSKS--LNGIY 212
QY 291 LDTAKQITKYNKHAFSGGRDTIEEHROFGNCVDVSEFMYLTFP--LEDDKLEQIRK 348
DB 213 LADDAADTLRKXMSMTYDPEDHIRVEDPGKIEGN---WVFHYLDVFGREPDAQEIADMK 268
QY 349 DYTSGAMLTGELKALIEVLOPLIAEHQARKEVTDEI 386
DB 269 RYRGGLGDVTKRYLLEILERELGPIRRRIEPAKM 306

RESULT 4

US-09-425-666-2

Sequence 2, Application US/09425666

Patent No. 6416976

GENERAL INFORMATION:

APPLICANT: Gentry, Danile

APPLICANT: Greenwood, Claire

APPLICANT: Lawlor, Elizabeth

TITLE OF INVENTION: No. 6416976el trps

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSER: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09425,666

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/928,100

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Gimm, Edward R

REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P31624-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 341 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-425-666-2

Query Match 8.8%; Score 185.5; DB 4; Length 341;
Best Local Similarity 24.0%; Pred. No. 3.2e-12;
Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 84 KPEVLTGKSPSEAMHGHLPFTKMLQDVFNVLQMTDDEKY-LMKDLTLDDQAY 142
DB 3 KPILL-TGDRPTCK-LHIGHVGSLSKNR-----VLLQEBKXDMFVFLADQQL 49
QY 143 GDAVEN-----AKDIIACGFDINKTFIF--SDLDYMGSSGFYKXNV---KIOK 186
DB 50 THAKDPQITVESIGNVALDYLAAGLDPKSTIFIQSQIPELAEISMTYMLVSLARLR 109
QY 187 HVTENQVKIGFPTSDCIKISFPALQAA--PSFNSFPQIFRDRTDIQCILPCALDD 244
DB 110 NPTVKTETISOKGFGESIPFGFLVYPLAQADITAFKANY-----VPVGTQK 156
QY 245 PYFRMTRD-----VAPRIGYKPKALHSTFPALQGAQTMSASDPNSIF 290
DB 157 PMIEQIREIVRSFNNAVNCVIVPEGIYPENE--RAGRLPGIDG--NAKMSKS--LNGIY 212
QY 291 LDTAKQITKYNKHAFSGGRDTIEEHROFGNCVDVSEFMYLTFP--LEDDKLEQIRK 348
DB 213 LADDAADTLRKXMSMTYDPEDHIRVEDPGKIEGN---WVFHYLDVFGREPDAQEIADMK 268
QY 349 DYTSGAMLTGELKALIEVLOPLIAEHQARKEVTDEI 386
DB 269 RYRGGLGDVTKRYLLEILERELGPIRRRIEPAKM 306

RESULT 5

US-08-743-130A-39

Sequence 39, Application US/08743130A

Patent No. 5871987

GENERAL INFORMATION:

APPLICANT: Sassanfar, Mandana

APPLICANT: Gallant, Paul L.

APPLICANT: Shen, Xiaoyu

APPLICANT: Tao, Nianjun

APPLICANT: Tao, Jianshi

APPLICANT: Houman, Fariba

TITLE OF INVENTION: CANDIDA TYROSYL-tRNA SYNTHETASE

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESSES:

ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Millitia Drive

CITY: Lexington

STATE: Massachusetts

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/743,130A

FILING DATE: 01-NOV-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Brook Esq., David E.

REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: CPI95-12

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240

TELEFAX: (617) 861-9540

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 409 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-743-130A-39

Query Match 7.8%; Score 164.5; DB 2; Length 409;
Best Local Similarity 20.7%; Pred. No. 1e-09;
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

QY 76 VLDAYENK-KPFLYTGRRSSSEAMVHGLIPFI-----FTKWLQDV-----F 117
DB 27 IKVLEKENRPVXIYWGTAFTGKP-HCGYFVPMIKLAHFLKAGCEVTLADLHAFLDNM 85
QY 118 NPELVIVQMTDDEKYLMDLTLDOAYGDVAVENAKDIIACGFDINKTFIFSDLDYMGSSGF 177
DB 86 KALEVVKYRKYRKYEFVVKAILKLSINVPIERLKFVVGSSYQKGGDV--MDLFLKSNIV 142
QY 178 YKRVVK-----IQHVTFNQVKGIFGPTSDCIGKISFPALQAPSPNSFPQIFPRDRTD 232
DB 143 SQNDARAGADVQKQVANPLLSGLI-----YPLMQA-----IDEBHLG 180
QY 233 IQCLIPCAIDDPYFEMTRDVAPRIGYKPKALLHSTFFPALQAGQTMASASPNSIFLT 292
DB 181 VDAQFG-GVDQKRIFLAENLPSIGYKRAHLMNPMVPGI-GQGGKMSASDPNSKIDII 238
QY 293 DTKAKQIKTKYKNAFSGG--RDT-----IEHRQPGN 323
DB 239 EEPKVVKKVNSAYCAPGELKONGLIAFLIEYVIOPIAELKTGVEGAFKLDIDPEKYG- 297
QY 324 CDVDSFMYLTFLEBDDKLEQIRKDYSGAMLTGELKALI---EVLQPLIAEHQARR 379
DB 298 ---PLSY-----DSIEQLKADPVDGKLAPPLKSGVADKINELLAPRAEPRESS- 343
QY 380 KEVTDEIVKEFMTPRK 395
DB 344 -----EEFQVAQK 351

RESULT 6
US-08-743-130A-2
Sequence 2, Application US/08743130A
Patent No. 5871987
GENERAL INFORMATION:
APPLICANT: Sagesanfar, Mandana
APPLICANT: Gallant, Paul L.
APPLICANT: Shen, Xiaoyu
APPLICANT: Tao, Mianjun
APPLICANT: Tao, Jianhui
APPLICANT: Houman, Fariba
TITLE OF INVENTION: CANDIDA TYROSYL-ERNA SYNTHETASE
TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743,130A
FILING DATE: 01-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook Esq., David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CP195-12
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-743-130A-2

Query Match 7.7%; Score 162.5; DB 2; Length 409;
Best Local Similarity 20.7%; Pred. No. 1.7e-09;
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

QY 76 VLDAYENK-KPFLYTGRRSSSEAMVHGLIPFI-----FTKWLQDV-----F 117
DB 27 IKVLEKENRPVXIYWGTAFTGKP-HCGYFVPMIKLAHFLKAGCEVTLADLHAFLDNM 85
QY 118 NPELVIVQMTDDEKYLMDLTLDOAYGDVAVENAKDIIACGFDINKTFIFSDLDYMGSSGF 177
DB 86 KALEVVKYRKYRKYEFVVKAILKLSINVPIERLKFVVGSSYQKGGDV--MDLFLKSNIV 142
QY 178 YKRVVK-----IQHVTFNQVKGIFGPTSDCIGKISFPALQAPSPNSFPQIFPRDRTD 232
DB 143 SQNDARAGADVQKQVANPLLSGLI-----YPLMQA-----IDEBHLG 180
QY 233 IQCLIPCAIDDPYFEMTRDVAPRIGYKPKALLHSTFFPALQAGQTMASASPNSIFLT 292
DB 181 VDAQFG-GVDQKRIFLAENLPSIGYKRAHLMNPMVPGI-GQGGKMSASDPNSKIDII 238
QY 293 DTKAKQIKTKYKNAFSGG--RDT-----IEHRQPGN 323
DB 239 EEPKVVKKVNSAYCAPGELKONGLIAFLIEYVIOPIAELKTGVEGAFKLDIDPEKYG- 297
QY 324 CDVDSFMYLTFLEBDDKLEQIRKDYSGAMLTGELKALI---EVLQPLIAEHQARR 379
DB 298 ---PLSY-----DSIEQLKADPVDGKLAPPLKSGVADKINELLAPRAEPRESS- 343
QY 380 KEVTDEIVKEFMTPRK 395
DB 344 -----EEFQVAQK 351

RESULT 7
US-08-705-868-4
Sequence 4, Application US/08705868
Patent No. 5885798
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Coleman, Roger
APPLICANT: Au-Young, Janice
APPLICANT: Murty, Lynn E.
TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,868
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0117 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 388 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 1184699
 US-08-705-868-4

Query Match 6.3%; Score 132.5; DB 2; Length 388;
 Best Local Similarity 20.8%; Pred. No. 3.6e-06;
 Matches 80; Conservative 58; Mismatches 121; Indels 125; Gaps 20;

QY 91 GRGPS-SEAMHVGHLIPFTKMLQDVFNVPVLIQMTDE--KYLKQDLTLDO--AYGD 144
 DB 2 GDAPSPEEKH-----LITRNLOEVLGEKRLKELKRELYWGATTKGPHVAFV 54
 QY 145 AVENKADIIACGFDINKTFIFSDL-----DYMGSFGFYKNVVKIOKH--YTF 190
 DB 55 PMSKIADFLKAGEV--TILFADLHAYLDNMKAPWELERVSYEYENIKAMLESIGVPL 112
 QY 191 NOVKGIFG-----FTDSD-----CIGKISFPALIOA 216
 DB 113 EKLKFKIGTDYOLSKETYLIDVYRLSSVTOHDSKKAAGAVVQVHEPHLISGLLYPGLQ- 171
 QY 217 PSFNSFPQIFDRDRIQCLIPCAIDODPYFRMTDVAPRIGYPPALLHSTFFPALOGA 276
 DB 172 --LDEEYLVK-----DAQF--GGIDQKIFTFPAEKYLPALGYSKRVHLMNPMVGLTG- 220
 QY 277 QTKMSASDPNSSIFLTDIAKQIKTKVKNKHAFF-----SGGRDTIEH----- 317
 DB 221 -SKMSSSEESKIDLLDRKEDVKKLKK-AFCBPNGVENNGVLSPIKHYLPPLKSEFVIL 278
 QY 318 --RQFGNCDVDVSPMYLTFLEDDDKLEQIRKDYTSGAMLTGELKKALIEVLOPLIAH 375
 DB 279 RDEKMGKN-----KTYTAYVD-----LEKDPAAEVVHPGDLKNSVEVALNKLL--- 321
 QY 376 QARRKEVTDEIVKEWTFP--RKLS 397
 DB 322 -----DPIREKENTPALKKLA 337

RESULT 8

US-09-123-615-4
 Sequence 4, Application US/09123615
 Patent No. 6090377
 GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 APPLICANT: Coleman, Roger
 APPLICANT: Au-Young, Janice
 APPLICANT: Murty, Lynn E.
 TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: U.S.
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/123,615
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/705,868
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0117 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 388 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 1184699
 US-09-123-615-4

Query Match 6.3%; Score 132.5; DB 3; Length 388;
 Best Local Similarity 20.8%; Pred. No. 3.6e-06;
 Matches 80; Conservative 58; Mismatches 121; Indels 125; Gaps 20;

QY 91 GRGPS-SEAMHVGHLIPFTKMLQDVFNVPVLIQMTDE--KYLKQDLTLDO--AYGD 144
 DB 2 GDAPSPEEKH-----LITRNLOEVLGEKRLKELKRELYWGATTKGPHVAFV 54
 QY 145 AVENKADIIACGFDINKTFIFSDL-----DYMGSFGFYKNVVKIOKH--YTF 190
 DB 55 PMSKIADFLKAGEV--TILFADLHAYLDNMKAPWELERVSYEYENIKAMLESIGVPL 112
 QY 191 NOVKGIFG-----FTDSD-----CIGKISFPALIOA 216
 DB 113 EKLKFKIGTDYOLSKETYLIDVYRLSSVTOHDSKKAAGAVVQVHEPHLISGLLYPGLQ- 171
 QY 217 PSFNSFPQIFDRDRIQCLIPCAIDODPYFRMTDVAPRIGYPPALLHSTFFPALOGA 276
 DB 172 --LDEEYLVK-----DAQF--GGIDQKIFTFPAEKYLPALGYSKRVHLMNPMVGLTG- 220
 QY 277 QTKMSASDPNSSIFLTDIAKQIKTKVKNKHAFF-----SGGRDTIEH----- 317
 DB 221 -SKMSSSEESKIDLLDRKEDVKKLKK-AFCBPNGVENNGVLSPIKHYLPPLKSEFVIL 278
 QY 318 --RQFGNCDVDVSPMYLTFLEDDDKLEQIRKDYTSGAMLTGELKKALIEVLOPLIAH 375
 DB 279 RDEKMGKN-----KTYTAYVD-----LEKDPAAEVVHPGDLKNSVEVALNKLL--- 321
 QY 376 QARRKEVTDEIVKEWTFP--RKLS 397
 DB 322 -----DPIREKENTPALKKLA 337

RESULT 9

US-08-855-910-11
 Sequence 11, Application US/08855910
 Patent No. 6221640
 GENERAL INFORMATION:
 APPLICANT: Tao, Jianshi
 APPLICANT: Sassanfar, Mandana
 APPLICANT: Gallant, Paul L.
 APPLICANT: Shen, Xiaoyu
 APPLICANT: Avruich, Anthony S.
 APPLICANT: Yu, Russell V.
 APPLICANT: Nair, Shanjia
 TITLE OF INVENTION: ENTEROCOCCAL AMINOACYL-tRNA SYNTHETASE
 TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
 NUMBER OF SEQUENCES: 71
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 STREET: Two Militia Drive
 CITY: Lexington
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02173
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/855,910
 FILING DATE: 14-MAY-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Brook, David E.
 REGISTRATION NUMBER: 22,592
 REFERENCE/DOCKET NUMBER: CPI95-08
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781) 861-6240
 TELEFAX: (781) 861-9540
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 418 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-855-910-11

Query Match 5.4%; Score 113.5; DB 4; Length 418;
 Best Local Similarity 20.5%; Pred. No. 0.00056;
 Matches 82; Conservative 47; Mismatches 112; Indels 159; Gaps 18;

QY 88 LYGRCSSAMVGHILPITFKWLODVENVPLV-----123
 DB 33 LYGVDPDGDMSHIGLIPFMMKRFLAGHHPYLLGGGTIGDPSGRTTERVLTQME 92
 QY 124 -----OMTDEKYLW-KOLTLDOAYGDAVEN-----AKDIIA 154
 DB 93 AVQHNVDLSNOMKLEFGKAEVMTVNNYDLSLSLIDPLRDYGKFNVTMLAKDIIA 152
 QY 155 CGPDINKTIFPSDLDYMGMS-GEYKVVVKIQKVTNNVKG-----195
 DB 153 S--RLSGISFTFTYQILQSIDFY---TLHKKNILQIGADQMGNTAGLDLRKE 207
 QY 196 ----TGF-----DSDICIGISFPALQAPSFNSFP--QIFRDRDIOCLIFCAI 241
 DB 208 GPEAKVGLTIPMLKADGTRFGKTAGAIMLDPKKTSPEFYQFWMNOD-----258
 QY 242 DQDPYFMTDVAVRIGYPKPALHSTFPFALQAGOTKMSASDPNSIFLDTAKOIKTK 301
 DB 259 -----RDV---ITLK-----FTFLDKER-----IDALAEVKE 286
 QY 302 VNKIAPSGGDTIEHROFGNCDVDVSFWYLFLEDDDKLQRIKDYTSGAMLTGELK 361
 DB 287 PKR--EAQRLE-----VTRFVHDDALBEAOK--ISEALLFSGNIK 326
 QY 362 KALIEVLQ-----PLAEHQAARKEVTDEIVKFFMTPRK 395
 DB 327 DLTIEIEGLHEHVFV-EITKAKNIVDWLVDTIEPSK 365

RESULT 10
 US-09-352-990-28
 ; Sequence 28, Application US/09352990
 ; Patent No. 6255090
 ; GENERAL INFORMATION:
 ; APPLICANT: Famodu, Layo O.
 ; APPLICANT: Oroczo, Buddy
 ; APPLICANT: Rafalski, Antoni
 ; TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
 ; FILE REFERENCE: BB-1191

CURRENT APPLICATION NUMBER: US/09/352,990
 ; CURRENT FILING DATE: 1999-07-14
 ; EARLIER APPLICATION NUMBER: 60/092,866
 ; EARLIER FILING DATE: July 15, 1998
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 28
 ; LENGTH: 377
 ; TYPE: PRT
 ; ORGANISM: Synecocystis sp.
 US-09-352-990-28

Query Match 5.2%; Score 109; DB 4; Length 377;
 Best Local Similarity 21.0%; Pred. No. 0.0015;
 Matches 82; Conservative 49; Mismatches 116; Indels 144; Gaps 19;

QY 68 FSHRDMNOVLDAVENKKPFLYLTGKPSSEAMVGHILPITFKWLODVENVPLVQMTD 127
 DB 33 FAHRSTTAM-----DKPRIL-SGVQPTGN-LHLGNVLGAI-RSWVEQ-----QOHY 75
 QY 128 DEKYLKDL-----TLDOAYGDAVENAKDIIAGPDINKTIF-----165
 DB 76 DNEFCVVDLHAITVPHNPQTLAQ---DTLTALVLAGIDLOYSTIFVQSHVAHSELA 132
 QY 166 -----SDLYMGSSGFYKVVVKIQKVTNNVKGIFGPTDSDICIGISFPALQAPSF 219
 DB 133 WLNCVTPLNLMERMQFKEKAVKQGENVS-----VGLDYPVLMAA-----174
 QY 220 SNSFPQIFRDRDIOCL-----IPCAIDQPYRMTD-----ARIGPKRA 263
 DB 175 -----DILLYDADKVPVGEDOKOHELTREDIVIRINDKFGREDAPVLPKPEPL 222
 QY 264 L-LHSTFPFALQAGOTKMSASDPN--SIFLDTAKOIKTKVKNKHAFGSGARDTIEHRQF 320
 DB 223 IRKGGARVMSLADGTGKMSKDESELSRLINLDPDMIKKTKK-----266
 QY 321 GGNCDVVSFWYLFLEDDDKLQRIKDYTSGAMLTGELKAL-----364
 DB 267 ---CKTDPO-RGLMF---DDPERECHNLLTYTLISNOTKRAVQECAMGWCQPKPL 319
 QY 365 -----IEVLOPLAEHQAARKEVTDEIVK 388
 DB 320 TETAIALPEIQKVAEILLADRGEL-DRITIQ 349

RESULT 11
 US-08-923-867-2
 ; Sequence 2, Application US/08923867
 ; Patent No. 5851809
 ; GENERAL INFORMATION:
 ; APPLICANT: Lawlor, Elizabeth
 ; TITLE OF INVENTION: NOVEL TRYPTOPHANYL tRNA SYNTHETASE
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: SmithKline Beecham Corporation
 ; STREET: 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19406-0939
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/923,867
 ; FILING DATE: 04-SEP-1997
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 9619072.3
 ; FILING DATE: 12-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:

NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31624
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-923-867-2

Query Match 5.1%; Score 108; DB 2; Length 197;
Best Local Similarity 25.3%; Pred. No. 0.00067;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 237 IPCAIDDDPYFRMTD-----VAPRIGYPRPALHSTFPALOGAQTMSA 282
DB 5 VEVGTDDQKPMIEQTRIVRSFNNAYNCVLAPEEGIYPENE--RAGRLEGLDG-NAKMSK 61

QY 283 SDPNSSIFLTPTAKQIKTKVKNKHAFFSGRDTIEHRQFGNCDVDVSFMYLTFF--LEDD 340
DB 62 S-LNNGIYIADDDADTLRKVKVMSMYTDPDHIRVEDPKIEGN---MVFRYLDVFGRPEDA 116

QY 341 DLEQIRKDYTSGAMLTGELKKALIEVLOPLAEHQARKKEVTDEI 386
DB 117 QEIAIMKERYQRGGLGDVTKRYLLEILERELGPFRERRIEPAKDM 162

RESULT 12
US-08-928-100-4
Sequence 4, Application US/08928100
Patent No. 6046174
GENERAL INFORMATION:
APPLICANT: Gentry, Danile
APPLICANT: Greenwood, Claire
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 6046174e1 trps
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,100
FILING DATE: 12-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9619072.3
FILING DATE: 12-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31624-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-928-100-4

Query Match 5.1%; Score 108; DB 3; Length 197;
Best Local Similarity 25.3%; Pred. No. 0.00067;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 237 IPCAIDDDPYFRMTD-----VAPRIGYPRPALHSTFPALOGAQTMSA 282
DB 5 VEVGTDDQKPMIEQTRIVRSFNNAYNCVLAPEEGIYPENE--RAGRLEGLDG-NAKMSK 61

QY 283 SDPNSSIFLTPTAKQIKTKVKNKHAFFSGRDTIEHRQFGNCDVDVSFMYLTFF--LEDD 340
DB 62 S-LNNGIYIADDDADTLRKVKVMSMYTDPDHIRVEDPKIEGN---MVFRYLDVFGRPEDA 116

QY 341 DLEQIRKDYTSGAMLTGELKKALIEVLOPLAEHQARKKEVTDEI 386
DB 117 QEIAIMKERYQRGGLGDVTKRYLLEILERELGPFRERRIEPAKDM 162

RESULT 13
US-09-183-134-2
Sequence 2, Application US/09183134
Patent No. 6165759
GENERAL INFORMATION:
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: NOVEL TRYPTOPHANYL tRNA
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/183,134
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/923,867
FILING DATE: 04-SEP-1997
APPLICATION NUMBER: 9619072.3
FILING DATE: 12-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31624
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-183-134-2

Query Match 5.1%; Score 108; DB 4; Length 197;
Best Local Similarity 25.3%; Pred. No. 0.00067;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

```

QY      237  IPCALDOPBYFRMTD-----VAPRIQYPKPALLHSIFPALQOAGTRMSA 282
Dh      5  VPGVDQPKMEIQTEIREIYRSPNNAINCDVLVEBEGYIPENE--RAORLFGJDS-NAKMSK 61
QY      283  SDPNSSIFLTPAKOIKTKVKNHAFSGAGDRJTIEHNRQFGANCVDVDSFWMLTFP--LEDD 340
Dh      62  S-LNNGIYIADDAADLTLRKKKVMYMDPDRIAREDDPKIEGN---MVFRHLDVFGRPEDA 116
QY      341  DKLEQIRDYDTSGAMLTGELKALIEVLQPLAIEHQARRKEVYTDI 386
Dh      117  QEIAIMKERYQGGGLGADVTKTKRYLLEITREBELGPIHERRIRIEPAKOM 162

```

RESULT 14
US-09-492-581-4
; Sequence 4, Application US/09492581
; Patent No. 5345400

APPLICANT: Gentry, Danile
 APPLICANT: Greenwood, Claire
 APPLICANT: Lawlor, Elizabeth
 TITLE OF INVENTION: No. 6346409e1 trps
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham Corporation
 STREET: 709 Swedeland Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19406-0939
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/492,581
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,100
 FILING DATE: 12-SEP-1997
 APPLICATION NUMBER: 9619072.3
 FILING DATE: 12-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Gimmil, Edward R.
 REGISTRATION NUMBER: 38,891
 REPRESENTATION/DOCKET NUMBER: P31624-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090
 TELETYPE:
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 197 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein

Query Match 5.1%; Score 108; DB 4; Length 197;
Best Local Similarity 25.3%; Pred No. 0.00067;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6

Dy 237 IPCADIDDPYFEMTDD-----VARIQYPKALHSITFPALQAQMSA 282
Db 5 VPGVDQDKMIEQTREIYVSFNNAVCNVLVEPEGIYPEHE--RAGRLLPGLDG-NAMKSK 61

Oy 283 SDPNSSIFLTTPAKQIKKTUVNKAFSGGSDTIENHQFGNCDCVDVSFWLITFF--LEDD 340
Db 62 S-LNNGIYIADDADTLTKRKVMKSYTDPDHIIVADPGCIGIRSN---WVFHLVDVFGRPEDA 116

```

QY      341 DKLEGRKDYTSGLMTLTKKALIEVLQPLTAHQARKKEVTDEI 386
      :: :: | | : : | : | : : : | | : :
DB      117 QEIDMKERYQRGGLGVKTRYLLEILEREIGPIRERRIEFAKDM 1622

```

RESULT 15
US-09-425-666-4
; Sequence 4, Application US/09425666
; Patent No. 6416076

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

?      FILING DATE:
?
?      ATTORNEY/AGENT INFORMATION:
?      NAME: Gimmil, Edward R
?      REGISTRATION NUMBER: 38,891
?      REFERENCE/DOCKET NUMBER: P31624-1
?
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: 610-270-4478
?      TELEFAX: 610-270-5090
?
?      TELEX:
?
?      INFORMATION FOR SEQ. ID NO.: 4:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 197 amino acids
?      TYPE: amino acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?
?      MOLECULE TYPE: protein
US-09-425-666-4

```

Query Match 5.1%; Score 108; DB 4; Length 197;
Best Local Similarity 25.3%; Pred. No. 0.00067;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6.

```

OY 337 1PCADIDPOYFRMRD-----VA9RGYKFPALSHSTFFPALQAGQKSA 282
Db 5  VPVGQDKPMTQE9FREIRVRSFNNAINCULVEPEGIYENE--RAGRFLUGLGG-NAKSK 61
OY 283 SDPNSIPLTQAQIKTKYKNAHAFSGGRDITEE9ROGNCDDVVSFMYLTFF--LEDD 340
Db 62 S-LNNGIYLTADADPLTRKCMWSMTDPDPIHINVEDGKIEGN---MVFIYLDVFG9PEBA 116
OY 341 DKLEQIRKDYTSGAMLTGELKKALIEVLOPLAEHQARKVETDEI 386
Db 117 QEIDMKERYOGGGIGDVKTRKYLELIERELGPIRERITEPAKOM 162

```

Search completed: July 10, 2003, 12:32:51
Job time : 13.1515 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:24:54 ; Search time 32.08 Seconds
(without alignments)
1665.633 Million cell updates/sec

Title: US-09-813-718-10_COPY_71_471

Perfect score: 2116
Sequence: 1 SNHGPDTEAEDFVDPMTV.....VTDEIVKEFMTPRKLSDFQ 401

Scoring table: BLOSUM62

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

A_Geneseq_1011002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2116	100.0	415	22	AA847617	Human superinini Ty
2	2116	100.0	415	23	AA813493	Human superinini Ty
3	2116	100.0	437	22	AA847616	Human mini TyRS.
4	2116	100.0	437	23	AA813492	Human mini tryptop
5	2116	100.0	484	22	AA847615	Human full-length
6	2116	100.0	484	21	AA813491	Human-tryptophany
7	2101	99.3	475	21	AA858230	Lung cancer associ
8	2096	99.1	471	20	AA805372	Human HCMV inducib
9	1988	94.0	382	22	AA847618	Human inactive TyR
10	1988	94.0	392	23	AA813494	Human inactive try

11	1370.5	64.8	430	22	ABB6766021
12	1370.5	64.8	430	22	ABB6766021
13	1304	61.6	402	21	AAAG326989
14	1304	61.6	426	21	AAAG326989
15	1218.5	57.6	424	22	AAAB669397
16	968	45.7	292	21	AAAG583177
17	8831	39.3	173	21	AAAG583177
18	803	37.9	385	22	AAAB964009
19	328	15.5	85	23	AAE13515
20	226.5	10.7	142	21	AAAB682191
21	202.5	9.6	341	23	AAEP296694
22	185.5	8.8	341	19	AAAB584422
23	185.5	8.8	341	22	AAU3785151
24	185.5	8.6	341	23	AABP269696
25	174.5	8.2	401	23	AAAB593248
26	164.5	7.8	409	20	AAAW942448
27	162.5	7.7	409	20	AAW942448
28	159.5	7.5	379	20	AAAY376232
29	159.5	7.5	378	20	AAAY376232
30	152	7.2	344	20	AAAY334516
31	147	6.9	46	23	AAE13515
32	146	6.9	339	22	AAU3830233
33	138.5	6.5	140	21	AAAG228293
34	134.5	6.4	525	22	AAAB607545
35	134	6.3	372	22	AAAB476131
36	134	6.3	372	22	AAE13490
37	132.5	6.3	372	22	AAAB476131
38	132.5	6.3	372	23	AAE1349888
39	132.5	6.3	536	22	AAAB476131
40	132.5	6.3	536	22	AAE1349888
41	128	6.0	334	22	AAU3544777
42	126.5	6.0	419	22	AAAB454171
43	123	5.8	347	21	AAAG319030
44	122.5	5.8	346	21	AAU946161
45	119	5.6	415	23	AAAG791699

ALIGNMENTS

RESULT 1	
AA047617	
ID	AA047617 standard; Protein; 415 AA.
XX	
AC	AA047617;
XX	
DT	07-JAN-2002 (first entry)
XX	
DE	Human supermini TrpRS.
XX	
KW	Tyrosyl-tRNA synthetase; TyRS; Rosemann fold nucleotide binding domain
KW	vascular endothelial cell function; burn; plastic surgery; abdomen;
KW	polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
KW	angiogenesis; graft; myocardial infarction; solid tumour; wound healing
KW	dermal ulcer; diabetic ulcer; endothelialization;
KW	tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.
XX	
OS	Homo sapiens.
XX	
PN	WO200174841-A1.
XX	
PD	11-OCT-2001.
XX	
PF	21-MAR-2001; 2001WO-US08966.
XX	
PR	31-MAR-2000; 2000US-193471P.
XX	
PA	(SCRI) SCRIPPS RES INST.
XX	
PI	Schimmel P, Wakasugi K;
XX	
WP	WP; 2001-626377/72.
DR	N-PSDB; AA043604.

XX New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
 PT vascular endothelial function, in particular for regulating
 PT angiogenesis, tumor metastasis and treating myocardial infarction
 PS Disclosure; Page 129-30; 150pp; English.

CC The sequences given in AAB47615-18 show full length and truncated
 CC versions of tyrophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
 CC the invention comprises a Rossmann fold nucleotide binding domain; and
 CC is capable of regulating vascular endothelial cell function. It is of
 CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of
 CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated
 CC TrpRS is useful for regulating angiogenesis, tumor metastasis,
 CC enhancing angiogenesis to a graft, treating myocardial infarction,
 CC solid tumor, and a condition that would benefit from increased or
 CC decreased angiogenesis in a mammal, in particular humans. It is also
 CC useful in diagnosis and as a wound healing agent for treating wounds
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
 CC plastic surgery where reconstruction is required following a burn or
 CC for cosmetic purposes. It is particularly useful in the treatment of
 CC abdominal wounds where there is high risk of infection. Truncated TrpRS
 CC promotes endothelialization in vascular graft surgery and is used in
 CC conjunction with angiography to administer the angiogenic RNA
 CC synthetase polypeptide or polynucleotides directly to the lumen and
 CC wall of the blood vessel.

XX Sequence 415 AA;

Query Match 100.0%; Score 2116; DB 22; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1.6e-209;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDTEAEEDFVDVPTVQSSAKGIDYDKLIVRGSSKIDKELINRERATGQRP 60
 DB 2 SNHGPDTEAEEDFVDVPTVQSSAKGIDYDKLIVRGSSKIDKELINRERATGQRP 61
 QY 61 FLRGIFESHRDNNQVLDAYENKPFYLYTGRSPSEAMVGHILPIFTKMLQDVFNVP 120
 DB 62 FLRGIFESHRDNNQVLDAYENKPFYLYTGRSPSEAMVGHILPIFTKMLQDVFNVP 121
 QY 121 LVIQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDMGSSGFYKN 180
 DB 122 LVIQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDMGSSGFYKN 181
 QY 181 VKIKQKVTNOKYKIGFSDSCIGKISPPALAAFSFNSPFOIRDRDTIOCLIPCA 240
 DB 182 VKIKQKVTNOKYKIGFSDSCIGKISPPALAAFSFNSPFOIRDRDTIOCLIPCA 241
 QY 241 IDDDPYFRMRDVAIRIGYKPKPALIHSTFPFALQAGQTKMSASDPNSIFLTDAKQIKT 300
 DB 242 IDDDPYFRMRDVAIRIGYKPKPALIHSTFPFALQAGQTKMSASDPNSIFLTDAKQIKT 301
 QY 301 KVKNAFSGGRDTIEHRQFGNCVDVSPMYLTFPLEDDDKLEQIRKDYTSQAMLTGEL 360
 DB 302 KVKNAFSGGRDTIEHRQFGNCVDVSPMYLTFPLEDDDKLEQIRKDYTSQAMLTGEL 361
 QY 361 KKALIEVLOPLIAHQARKKEVTDIVKEPMPKLSFDQ 401
 DB 362 KKALIEVLOPLIAHQARKKEVTDIVKEPMPKLSFDQ 402

RESULT 2
 AAE13493
 ID AAE13493 standard; Protein; 415 AA.
 AC AAE13493;
 XX
 DT 12-FEB-2002 (first entry)
 XX
 DE Human supermini tyrophanyl t-RNA synthetase in pET20B.
 XX
 KW Human; tyrophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;

KW TyRS; vascular endothelial cell function; angiogenesis; wound healing;
 KW re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;
 KW diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;
 KW angiography; gene therapy; tumour; inflammation; vascular permeability;
 KW rheumatoid arthritis; psoriasis; diabetic retinopathy.

OS Homo sapiens.
 XX
 PN WO200175078-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 21-MAR-2001; 2001WO-US08975.
 XX
 PR 31-MAR-2000; 2000US-193471P.
 XX
 PA (SCRI) SCRIPPS RES INST.
 XX
 PI Schimmel P, Makaeygi K;
 XX
 DR WPI; 2002-010784/01.
 DR N-PSDB; AAD22484.
 XX
 PT Novel truncated tyrophanyl-tRNA synthetase polypeptides capable of
 PT regulating vascular endothelial cell function, preferably angiogenesis,
 PT is useful for treating solid tumor or suppressing tumor metastasis in
 PT mammal -

XX Example 1; Page 129-130; 149pp; English.

XX The patent discloses human aminoacyl tRNA synthetases, particularly
 CC truncated tyrophanyl-tRNA synthetases (TrpRS) comprising a Rossmann
 CC fold nucleotide binding domain and polynucleotides encoding them. The
 CC invention also relates to tyrosyl t-RNA synthetases (TyRS). TrpRS
 CC sequences are useful for regulating vascular endothelial cell function,
 CC preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound
 CC healing agents for re-vascularising damaged tissues. They are useful for
 CC treating full-thickness wounds (e.g. dermal ulcers), including pressure
 CC sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS
 CC sequences can also be used in plastic surgery when reconstruction is
 CC required following a burn, other trauma, or even for cosmetic purposes.
 CC Angiogenic TrpRS is also used in association with surgery and following
 CC the repair of cuts, for promoting endothelialisation in vascular graft
 CC surgery and for repairing the damage of myocardial infarction and in
 CC conjunction with coronary bypass surgery by stimulating the growth of
 CC transplanted tissue. TrpRS is also used in conjunction with angiography.
 CC TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in
 CC immunoassays to detect the presence of tumours. They are also useful
 CC for blocking endogenous angiogenic activity and retard the growth of
 CC solid tumours. These antibodies may also be used to treat inflammation
 CC caused by increased vascular permeability. Inhibiting the activity of
 CC TrpRS by antisense technology is useful for preventing further growth
 CC or even regress solid tumours, and for treating rheumatoid arthritis,
 CC psoriasis, diabetic retinopathy, all of which are characterised by
 CC abnormal angiogenesis. The present sequence is human truncated
 CC tyrophanyl t-RNA synthetase (supermini TrpRS; residues 71-471 of
 CC full-length TrpRS protein) in pET20B.

XX Sequence 415 AA;

Query Match 100.0%; Score 2116; DB 23; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1.6e-209;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDTEAEEDFVDVPTVQSSAKGIDYDKLIVRGSSKIDKELINRERATGQRP 60
 DB 2 SNHGPDTEAEEDFVDVPTVQSSAKGIDYDKLIVRGSSKIDKELINRERATGQRP 61
 QY 61 FLRGIFESHRDNNQVLDAYENKPFYLYTGRSPSEAMVGHILPIFTKMLQDVFNVP 120
 DB 62 FLRGIFESHRDNNQVLDAYENKPFYLYTGRSPSEAMVGHILPIFTKMLQDVFNVP 121
 QY 121 LVIQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDMGSSGFYKN 180

```

Db 122 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPDINKTIFSDLDYMGSSGFYKN 181
Qy 181 VKIKQHTFNVQKGI FGTSDSCIGKISFPALIOAAPSPNSFPQIFRDRDIOCLIPCA 240
Db 182 VKIKQHTFNVQKGI FGTSDSCIGKISFPALIOAAPSPNSFPQIFRDRDIOCLIPCA 241
Qy 241 IDDDPYFMTTRDVARIGYKPKPALHSTFPPALOGAQTMSADPNSSIFLDTAKQIKT 300
Db 242 IDDDPYFMTTRDVARIGYKPKPALHSTFPPALOGAQTMSADPNSSIFLDTAKQIKT 301
Qy 301 KVNKAHFGSGRDTIEHRQFGNCVDVSPMYLTFLEDDDKLEQIRKDYTSGAMLTGEL 360
Db 302 KVNKAHFGSGRDTIEHRQFGNCVDVSPMYLTFLEDDDKLEQIRKDYTSGAMLTGEL 361
Qy 361 KKALEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401
Db 362 KKALEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 402

```

RESULT 3

AA047616 standard; Protein; 437 AA.

AA047616; (first entry)

Human mini TrpRS.

Tyrosyl-tRNA synthetase; TrpRS; Rossmann fold nucleotide binding domain; vascular endothelial cell function; burn; plastic surgery; abdomen; polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis; dermal ulcer; diabetic ulcer; myocardial infarction; solid tumour; wound healing; tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.

Homo sapiens.

WO200174841-A1.

11-OCT-2001.

21-MAR-2001; 2001WO-US08966.

31-MAR-2000; 2000US-193471P.

(SCRI) SCRIPPS RES INST.

Schimmel P, Wakaugui K;

WPI; 2001-626377/72.

N-PSDB; AA043603.

New human truncated tyrosyl-tRNA synthetase polypeptide for regulating vascular endothelial function, in particular for regulating angiogenesis, tumor metastasis and treating myocardial infarction - Disclosure; Page 123-24; 150pp; English.

The sequences given in AA047615-18 show full length and truncated versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of the invention comprises a Rossmann fold nucleotide binding domain, and is capable of regulating vascular endothelial cell function. It is of approx. 40 kilo Dalton molecular weight and is produced by cleavage of full length TrpRS with polymorphonuclear leucocyte elastase. Truncated TrpRS is useful for regulating angiogenesis, tumor metastasis, enhancing angiogenesis to a graft, treating myocardial infarction, solid tumor, and a condition that would benefit from increased or decreased angiogenesis in a mammal, in particular humans. It is also useful in diagnosis and as a wound healing agent for treating wounds such as dermal ulcers, diabetic ulcers, burns and injuries and in plastic surgery when reconstruction is required following a burn or

CC for cosmetic purposes. It is particularly useful in the treatment of CC abdominal wounds where there is high risk of infection. Truncated TrpRS CC promotes endothelialization in vascular graft surgery and is used in CC conjunction with angiography to administer the angiogenic RNA CC synthetase polypeptides or polynucleotides directly to the lumen and CC wall of the blood vessel.

SO Sequence 437 AA;

Query Match 100.0%; Score 2116; DB 22; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.8e-209;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 SNHGDAEAEDPDPMTVQTSKAGIDYDLIVRFGSSKIDKLINRIEATGQRPNN 60
Db 24 SNHGDAEAEDPDPMTVQTSKAGIDYDLIVRFGSSKIDKLINRIEATGQRPNN 83
Qy 61 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSPSEAMHVGHLPIFTKMLQDVNVP 120
Db 84 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSPSEAMHVGHLPIFTKMLQDVNVP 143
Qy 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPDINKTIFSDLDYMGSSGFYKN 180
Db 144 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPDINKTIFSDLDYMGSSGFYKN 203
Qy 181 VKIKQHTFNVQKGI FGTSDSCIGKISFPALIOAAPSPNSFPQIFRDRDIOCLIPCA 240
Db 204 VKIKQHTFNVQKGI FGTSDSCIGKISFPALIOAAPSPNSFPQIFRDRDIOCLIPCA 263
Qy 241 IDDDPYFMTTRDVARIGYKPKPALHSTFPPALOGAQTMSADPNSSIFLDTAKQIKT 300
Db 264 IDDDPYFMTTRDVARIGYKPKPALHSTFPPALOGAQTMSADPNSSIFLDTAKQIKT 323
Qy 301 KVNKAHFGSGRDTIEHRQFGNCVDVSPMYLTFLEDDDKLEQIRKDYTSGAMLTGEL 360
Db 324 KVNKAHFGSGRDTIEHRQFGNCVDVSPMYLTFLEDDDKLEQIRKDYTSGAMLTGEL 383
Qy 361 KKALEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401
Db 384 KKALEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 424

```

RESULT 4

AAE13492 standard; Protein; 437 AA.

AAE13492;

12-FEB-2002 (first entry)

Human mini tryptophanyl t-RNA synthetase in pET20B.

Human; tryptophanyl-tRNA synthetase; TrpRS; tyrosyl-tRNA synthetase; TrpRS; vascular endothelial cell function; angiogenesis; wound healing; re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury; diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction; angiography; gene therapy; tumour; inflammation; vascular permeability; rheumatoid arthritis; psoriasis; diabetic retinopathy.

Homo sapiens.

WO200175078-A1.

11-OCT-2001.

21-MAR-2001; 2001WO-US08975.

31-MAR-2000; 2000US-193471P.

(SCRI) SCRIPPS RES INST.

Schimmel P, Wakaugui K;

XX

DR WPI: 2002-010784/01.
DR N-PSDB; AAD22483.

PT Novel truncated tryptophanyl-tRNA synthetase polypeptides capable of
PT regulating vascular endothelial cell function, preferably angiogenesis,
PT is useful for treating solid tumor or suppressing tumor metastasis in
PT mammal -

PS Example 1; Page 123-124; 149pp; English.

XX The patent discloses human aminoacyl tRNA synthetases, particularly
XX truncated tryptophanyl-tRNA synthetases (TrpRS) comprising a Rossmann
XX fold nucleotide binding domain and polynucleotides encoding them. The
XX invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS
XX sequences are useful for regulating vascular endothelial cell function,
XX preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound
XX healing agents for re-vascularizing damaged tissues. They are useful for
XX treating full-thickness wounds (e.g. dermal ulcers, including pressure
XX sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS
XX sequences can also be used in plastic surgery when reconstruction is
XX required following a burn, other trauma, or even for cosmetic purposes.
XX Angiogenic TrpRS is also used in association with surgery and following
XX the repair of cuts, for promoting endothelialisation in vascular graft
XX surgery and for repairing the damage of myocardial infarction and in
XX conjunction with coronary bypass surgery by stimulating the growth of
XX transplanted tissue. TrpRS is also used in conjunction with angiography.
XX TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in
XX immunoassays to detect the presence of tumours. They are also useful
XX for blocking endogenous angiogenic activity and retard the growth of
XX solid tumours. These antibodies may also be used to treat inflammation
XX caused by increased vascular permeability. Inhibiting the activity of
XX TrpRS by antisense technology is useful for preventing further growth
XX or even regress solid tumours, and for treating rheumatoid arthritis,
XX psoriasis, diabetic retinopathy, all of which are characterised by
XX abnormal angiogenesis. The present sequence is human truncated
XX tryptophanyl t-RNA synthetase (mini TrpRS; residues 48-471 of
XX full-length TrpRS protein) protein in PET20B.

XX Sequence 437 AA;

Query Match 100.0%; Score 2116; DB 23; Length 437;

Best Local Similarity 100.0%; Pred. No. 1.8e-209; Mismatches 0; Indels 0; Gaps 0;

Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDTEABEDPVDPMVTQTSAAKIDYDKLIVFGSSKIDKELINRIERATGQRPH 60
DB 24 SNHGPDTEABEDPVDPMVTQTSAAKIDYDKLIVFGSSKIDKELINRIERATGQRPH 83
QY 61 FLRGIFFSHRDMNQVLDAVENKKPPYLYTGRGSSSEAMVGHILPFIPTKWLQDVENV 120
DB 84 FLRGIFFSHRDMNQVLDAVENKKPPYLYTGRGSSSEAMVGHILPFIPTKWLQDVENV 143
QY 121 LVITQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSLDYMGSSEGYKN 180
DB 144 LVITQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSLDYMGSSEGYKN 203
QY 181 VVKTIQKHTVENOVKIGFPTSDGIGKISPPAIOAASFNSPQIRDRDTEIQCLIPCA 240
DB 204 VVKTIQKHTVENOVKIGFPTSDGIGKISPPAIOAASFNSPQIRDRDTEIQCLIPCA 263
QY 241 IDDDPYFRMTRDVAIRIGYKPPALHSTFPALQAGTQKMSASDPNSSIFLDTAKOIKT 300
DB 264 IDDDPYFRMTRDVAIRIGYKPPALHSTFPALQAGTQKMSASDPNSSIFLDTAKOIKT 323
QY 301 KVNKHAFSGGRDTEEHROFGKNCDDVVSFMYLTLPLEDDDKLEQIRKDYTSGAMLTGEL 360
DB 324 KVNKHAFSGGRDTEEHROFGKNCDDVVSFMYLTLPLEDDDKLEQIRKDYTSGAMLTGEL 383
QY 361 KKALIEVLQPLIAEHQARRKEVDIEYKEPMPKLSFDFQ 401
DB 384 KKALIEVLQPLIAEHQARRKEVDIEYKEPMPKLSFDFQ 424

RESULT 5

ID AAB47615 standard; Protein; 484 AA.

XX AAB47615;

AC 07-JAN-2002 (first entry)

DE Human full-length TrpRS.

XX Tyrosyl-tRNA synthetase; TyrRS; Rossmann fold nucleotide binding domain;
XX vascular endothelial cell function; burn; plastic surgery; abdomen;
XX polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
XX angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
XX dermal ulcer; diabetic ulcer; endothelialization;
XX tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.

XX Homo sapiens.

PN W0200174841-A1.

PD 11-OCT-2001.

XX 21-MAR-2001; 2001WO-US08966.

XX 31-MAR-2000; 2000US-193471P.

XX (SCRI) SCRIPPS RES INST.

XX Schimmel P, Wakaugui K;

XX WPI; 2001-626377/72.

DR N-PSDB; AAB43602.

PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating

PT vascular endothelial function, in particular for regulating

PT angiogenesis, tumor metastasis and treating myocardial infarction

XX Disclosure; Page 117-19; 150pp; English.

XX The sequences given in AAB47615-18 show full length and truncated
XX versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
XX the invention comprises a Rossmann fold nucleotide binding domain, and
XX is capable of regulating vascular endothelial cell function. It is of
XX approx. 40 kilo Dalton molecular weight and is produced by cleavage of
XX full length TrpRS with polymorphonuclear leucocyte elastase. Truncated
XX TrpRS is useful for regulating angiogenesis, tumor metastasis,
XX enhancing angiogenesis to a graft, treating myocardial infarction,
XX solid tumor, and a condition that would benefit from increased or
XX decreased angiogenesis in a mammal, in particular humans. It is also
XX useful in diagnosis and as a wound healing agent for treating wounds
XX such as dermal ulcers, diabetic ulcers, burns and injuries and in
XX plastic surgery when reconstruction is required following a burn or
XX for cosmetic purposes. It is particularly useful in the treatment of
XX abdominal wounds where there is high risk of infection. Truncated TrpRS
XX promotes endothelialization in vascular graft surgery and is used in
XX conjunction with angiography to administer the angiogenic tRNA
XX synthetase polypeptides or polynucleotides directly to the lumen and
XX wall of the blood vessel.

XX Sequence 484 AA;

Query Match 100.0%; Score 2116; DB 22; Length 484;

Best Local Similarity 100.0%; Pred. No. 2.1e-203; Mismatches 0; Indels 0; Gaps 0;

Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDTEABEDPVDPMVTQTSAAKIDYDKLIVFGSSKIDKELINRIERATGQRPH 60
DB 71 SNHGPDTEABEDPVDPMVTQTSAAKIDYDKLIVFGSSKIDKELINRIERATGQRPH 130
QY 61 FLRGIFFSHRDMNQVLDAVENKKPPYLYTGRGSSSEAMVGHILPFIPTKWLQDVENV 120
DB 131 FLRGIFFSHRDMNQVLDAVENKKPPYLYTGRGSSSEAMVGHILPFIPTKWLQDVENV 190

QY 121 LVITQMTDEKYLKMDLTDOAGDAVENAKDIIAGGPDINKTIFESDLDMGMSGGFYKN 180
 DB 191 LVITQMTDEKYLKMDLTDOAGDAVENAKDIIAGGPDINKTIFESDLDMGMSGGFYKN 250
 QY 181 VKIKQHTVFNOVKYGFEGFTDSDCIKISFPALQAPSPNSFPQIFRRTDIQCLIPCA 240
 DB 251 VKIKQHTVFNOVKYGFEGFTDSDCIKISFPALQAPSPNSFPQIFRRTDIQCLIPCA 310
 QY 241 IDDPYFMTRDVAPRIGYKPKALHSTFPALQAGQTMASADPNSSIFLDTAKQIKT 300
 DB 311 IDDPYFMTRDVAPRIGYKPKALHSTFPALQAGQTMASADPNSSIFLDTAKQIKT 370
 QY 301 KVNKAHFGSGRDTIEHRQFGGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360
 DB 371 KVNKAHFGSGRDTIEHRQFGGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 430
 QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFPDQ 401
 DB 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFPDQ 471

RESULT 6
 AAE13491
 ID AAE13491 standard; Protein; 484 AA.
 XX AAE13491;
 AC 12-FEB-2002 (first entry)
 DT 12-FEB-2002 (first entry)
 XX
 DE Human tryptophanyl t-RNA synthetase (TrpRS) in PET20B.
 XX
 KM Human; tryptophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;
 KM TyRS; vascular endothelial cell function; angiogenesis; wound healing;
 KM re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;
 KM diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;
 KM angiography; gene therapy; tumour; inflammation; vascular permeability;
 KM rheumatoid arthritis; psoriasis; diabetic retinopathy.
 XX
 OS Homo sapiens.
 XX
 PN WO200175078-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 21-MAR-2001; 2001WO-US08975.
 XX
 PR 31-MAR-2000; 2000US-193471P.
 XX
 PA (SCRI) SCRIPPS RES INST.
 XX
 PI Schimmel P, Wakaugui K;
 XX
 DR WPI; 2002-010784/01.
 XX
 DR N-PSDB; AAD22482.
 XX
 PT Novel truncated tryptophanyl-tRNA synthetase polypeptides capable of
 PT regulating vascular endothelial cell function, preferably angiogenesis,
 PT is useful for treating solid tumor or suppressing tumor metastasis in
 PT mammal -
 XX
 XX Example 1; Page 117-119; 149pp; English.

CC required following a burn, other trauma, or even for cosmetic purposes.
 CC Angiogenic TrpRS is also used in association with surgery and following
 CC the repair of cuts, for promoting endothelialisation in vascular graft
 CC surgery and for repairing the damage of myocardial infarction and in
 CC conjunction with coronary bypass surgery by stimulating the growth of
 CC transplanted tissue. TrpRS is also used in conjunction with angiography.
 CC TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in
 CC immunoassays to detect the presence of tumours. They are also useful
 CC for blocking endogenous angiogenic activity and retard the growth of
 CC solid tumours. These antibodies may also be used to treat inflammation
 CC caused by increased vascular permeability. Inhibiting the activity of
 CC TrpRS by antisense technology is useful for preventing further growth
 CC or even regress solid tumours, and for treating rheumatoid arthritis,
 CC psoriasis, diabetic retinopathy, all of which are characterised by
 CC abnormal angiogenesis. The present sequence is human tryptophanyl
 CC t-RNA synthetase (TrpRS) in PET20B.
 XX

SQ Sequence 484 AA.
 Query Match 100.0%; Score 2116; DB 23; Length 484;
 Best Local Similarity 100.0%; Pred. No. 2.1e-209;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFYDPWTVQTSAGKIDYDKLIYRFGSSKIDKELINRIERATGQRPHN 60
 DB 71 SNHGPDATAEEDFYDPWTVQTSAGKIDYDKLIYRFGSSKIDKELINRIERATGQRPHN 130
 QY 61 FLRRGIFFSHRDMNOVLDAYENKKPPYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 120
 DB 131 FLRRGIFFSHRDMNOVLDAYENKKPPYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 190
 QY 121 LVITQMTDEKYLKMDLTDOAGDAVENAKDIIAGGPDINKTIFESDLDMGMSGGFYKN 180
 DB 191 LVITQMTDEKYLKMDLTDOAGDAVENAKDIIAGGPDINKTIFESDLDMGMSGGFYKN 250
 QY 181 VKIKQHTVFNOVKYGFEGFTDSDCIKISFPALQAPSPNSFPQIFRRTDIQCLIPCA 240
 DB 251 VKIKQHTVFNOVKYGFEGFTDSDCIKISFPALQAPSPNSFPQIFRRTDIQCLIPCA 310
 QY 241 IDDPYFMTRDVAPRIGYKPKALHSTFPALQAGQTMASADPNSSIFLDTAKQIKT 300
 DB 311 IDDPYFMTRDVAPRIGYKPKALHSTFPALQAGQTMASADPNSSIFLDTAKQIKT 370
 QY 301 KVNKAHFGSGRDTIEHRQFGGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360
 DB 371 KVNKAHFGSGRDTIEHRQFGGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 430
 QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFPDQ 401
 DB 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFPDQ 471

RESULT 7
 AAB58220
 ID AAB58220 standard; Protein; 475 AA.
 XX AAB58220;
 AC 14-MAR-2001 (first entry)
 DT 14-MAR-2001 (first entry)
 XX
 DE Lung cancer associated polypeptide sequence SEQ ID 558.
 XX
 XX Human; lung cancer associated protein; neuroprotective; cytostatic;
 XX cardioactive; immunomodulatory; muscular active; vulnerary;
 XX gastroprotective; nephroprotective; anti-infective; gynecological;
 XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 XX proliferative disorder; wound healing; infectious disease.
 OS Homo sapiens.
 XX
 XX
 PN WO200055180-A2.
 XX
 PD 21-SEP-2000.


```

XX 08-MAR-2000; 2000MO-US05918.
XX PF 12-MAR-1999; 99US-0124270.
XX PR 12-MAR-1999; 99US-0124270.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX PI Ruben SM;
XX DR WPI; 2000-587514/55.
XX DR N-PSDB; AAF18096.
XX PT Lung cancer associated gene sequences, referred to as lung cancer
XX PT anti-gene, useful for treatment, prevention, and diagnosis of disorders
XX PT such as lung cancer -
XX PS Claim 11, Page 1052-1053; 1425pp; English.
XX CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
XX CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
XX CC associated proteins and polynucleotide sequences, their agonists, and
XX CC antagonists may have neuroprotective; cytostatic; cardioactive;
XX CC immunomodulatory; muscular; active general; vulnery; gastrointestinal
XX CC general; nephrotoxic; anti-infective; gynecological; or antibacterial
XX CC activity. The invention also includes antibodies specific for the
XX CC protein or polynucleotide sequences. The lung cancer associated
XX CC polynucleotide sequences may be used for detection of lung cancer,
XX CC chromosome identification, as chromosome markers, and for numerous other
XX CC diagnostic or research purposes. The proteins may be used to treat
XX CC disorders such as neural, immune, muscular, reproductive,
XX CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX CC disorders. The proteins may also be used in the treatment of wounds and
XX CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
XX CC peptide AAB58549 are used in the course of the invention for the
XX CC identification and characterisation of the polynucleotide and protein
XX CC sequences.
XX SQ Sequence 475 AA;
XX
XX Query Match 99.3%; Score 2101; DB 21; Length 475;
XX Best Local Similarity 99.5%; Pred. No. 7.1e-208;
XX Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 SNHGPDTEAEEDFVDPWTVQTSASAKGIDYDKLIVFGSSKIDKELINRERATGQRP 60
XX DB 75 SNHGPDTEAEEDFVDPWTVQTSASAKGIDYDKLIVFGSSKIDKELINRERATGQRP 134
XX
XX 61 FLRGGIFFSHRDMNQVLDAYENKRPFLYLTGRGSSSEAMHVGHLIPFTKMLQDVNVP 120
XX DB 135 FLRGGIFFSHRDMNQVLDAYENKRPFLYLTGRGSSSEAMHVGHLIPFTKMLQDVNVP 194
XX
XX 121 LVLTQMTDEKYLWKDILTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDMGSSGFFYN 180
XX DB 195 LVLTQMTDEKYLWKDILTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDMGSSGFFYN 254
XX
XX 181 VVKIQKHVTENQVKGIFGFTSDSCIGKISFPALQAAAPSFSNSFPQIFRRTDIOCLIPCA 240
XX DB 255 VVKIQKHVTENQVKGIFGFTSDSCIGKISFPALQAAAPSFSNSFPQIFRRTDIOCLIPCA 314
XX
XX 241 IDDDPYFRMTRDVAAPRIGYKPKPALHSTFPALQAGOTKMSASDPNSIFLTDIAKOIKT 300
XX DB 315 IDDDPYFRMTRDVAAPRIGYKPKPALHSTFPALQAGOTKMSASDPNSIFLTDIAKOIKT 374
XX
XX 301 KVNKHAFFSGGRDITTEHRQFGNCDDVVSFMYLTFPLEDDDKLEQIRKDYSGAMLTGEL 360
XX DB 375 KVNKHAFFSGGRDITTEHRQFGNCDDVVSFMYLTFPLEDDDKLEQIRKDYSGAMLTGEL 434
XX
XX 361 KKALIEVLQPLIAHQAARKEVTDIEIVKEFMTPKLSFDPQ 401
XX DB 435 KKALIEVLQPLIAHQAARKEVTDIEIVKEFMTPKLSFDPQ 475

```

```

RESULT 8
AAV05372
ID AAV05372 standard; Protein; 471 AA.
XX AC AAV05372;
XX AC AAV05372;
XX DT 30-JUN-1999 (first entry)
XX DE Human HCMV inducible gene protein, SEQ ID NO 12.
XX KW HCMV inducible gene; cigs; human; human cytomagalovirus; interferon;
XX KM anti-viral therapy; anti-HCMV therapy; detection; diagnosis;
XX KM drug screening.
XX OS Homo sapiens.
XX PN WO913075-A2.
XX PD 18-MAR-1999.
XX PF 08-SEP-1998; 98MO-US18638.
XX PR 22-SEP-1997; 97US-0059725.
XX PR 08-SEP-1997; 97US-0058180.
XX PA (UYPR-) UNIV PRINCETON.
XX PI Cong J, Schenk T, Zhu H;
XX DR WPI; 1999-243729/20.
XX DR N-PSDB; AAX33942.
XX
XX New isolated human genes
XX PS Claim 3; Page 112-114; 184pp; English.
XX CC This sequence is encoded by a human gene of the invention, and is induced
XX CC to express by both HCMV and interferon (IFN), designated HCMV-inducible
XX CC genes (cig or cigs). The invention also relates to genes that are
XX CC repressed in the presence of HCMV infection, designated HCMV-repressible
XX CC genes (crg or crgs). The products can be used to obtain agents which can
XX CC be used for anti-viral therapy, particularly anti-HCMV therapy. They can
XX CC also be used for the development of drugs that would allow for higher
XX CC dosage IFN treatments without the concomitant toxicity normally
XX CC associated with administering high levels of IFN. The products can also
XX CC be used for detection, diagnosis and drug screening.
XX SQ Sequence 471 AA;
XX
XX Query Match 99.1%; Score 2096; DB 20; Length 471;
XX Best Local Similarity 99.3%; Pred. No. 2.3e-207;
XX Matches 398; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 SNHGPDTEAEEDFVDPWTVQTSASAKGIDYDKLIVFGSSKIDKELINRERATGQRP 60
XX DB 71 SNHGPDTEAEEDFVDPWTVQTSASAKGIDYDKLIVFGSSKIDKELINRERATGQRP 130
XX
XX 61 FLRGGIFFSHRDMNQVLDAYENKRPFLYLTGRGSSSEAMHVGHLIPFTKMLQDVNVP 120
XX DB 131 FLRGGIFFSHRDMNQVLDAYENKRPFLYLTGRGSSSEAMHVGHLIPFTKMLQDVNVP 190
XX
XX 121 LVLTQMTDEKYLWKDILTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDMGSSGFFYN 180
XX DB 191 LVLTQMTDEKYLWKDILTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDMGSSGFFYN 250
XX
XX 181 VVKIQKHVTENQVKGIFGFTSDSCIGKISFPALQAAAPSFSNSFPQIFRRTDIOCLIPCA 240
XX DB 251 VVKIQKHVTENQVKGIFGFTSDSCIGKISFPALQAAAPSFSNSFPQIFRRTDIOCLIPCA 310
XX
XX 241 IDDDPYFRMTRDVAAPRIGYKPKPALHSTFPALQAGOTKMSASDPNSIFLTDIAKOIKT 300
XX DB 311 IDDDPYFRMTRDVAAPRIGYKPKPALHSTFPALQAGOTKMSASDPNSIFLTDIAKOIKT 370

```


QY 301 KVNKAHSGGRDTEIEHROFGNCVDVSMYLTFFLEDDDKLEQIRKDYTGAMLTGL 360
 DB 371 KVNKAHSGGRDTEIEHROFGNCVDVSMYLTFFLEDDDKLEQIRKDYTGAMLTGL 430
 QY 361 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSPDFQ 401
 DB 431 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSPDFQ 471

RESULT 9

AA047618
 ID AAB47618 standard; Protein; 392 AA.

AC AAB47618;
 DT 07-JAN-2002 (first entry)

DE Human inactive TrpRS.

KW Tyrosyl-tRNA synthetase; TrpRS; Rossmann fold nucleotide binding domain;
 KW vascular endothelial cell function; burn; plastic surgery; abdomen;
 KW polymorphonuclear leucocyte elastase; angiogenesis; tumor metastasis;
 KW angiogenesis; graft; myocardial infarction; solid tumor; wound healing;
 KW dermal ulcer; diabetic ulcer; endothelialization;
 KW tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.

OS Homo sapiens.

PN WO200174841-A1.

PD 11-OCT-2001.

PF 21-MAR-2001; 2001WO-US08966.

PR 31-MAR-2000; 2000US-193471P.

PA (SCRI) SCRIPPS RES INST.

PI Schimmel P, Makasugi K;

DR WPI; 2001-626377/72.

DR N-PSDB; AA043605.

PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
 PT vascular endothelial function, in particular for regulating
 PT angiogenesis, tumor metastasis and treating myocardial infarction
 PS Disclosure; Page 135-36; 150pp; English.

CC The sequences given in AAB47615-18 show full length and truncated
 CC versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
 CC the invention comprises a Rossmann fold nucleotide binding domain, and
 CC is capable of regulating vascular endothelial cell function. It is of
 CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of
 CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated
 CC TrpRS is useful for regulating angiogenesis, tumor metastasis,
 CC enhancing angiogenesis to a graft, treating myocardial infarction,
 CC solid tumor, and a condition that would benefit from increased or
 CC decreased angiogenesis in a mammal, in particular humans. It is also
 CC useful in diagnosis and as a wound healing agent for treating wounds
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
 CC plastic surgery when reconstruction is required following a burn or
 CC for cosmetic purposes. It is particularly useful in the treatment of
 CC abdominal wounds where there is high risk of infection. Truncated TrpRS
 CC promote endothelialization in vascular graft surgery and is used in
 CC conjunction with angiography to administer the angiogenic RNA
 CC synthetase polypeptides or polynucleotides directly to the lumen and
 CC wall of the blood vessel.

SO Sequence 392 AA;

Query Match 94.0%; Score 1988; DB 22; Length 392;
 Best Local Similarity 100.0%; Pred. No. 2.5e-196;

Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 SAKGIDYDKLIVRFSSSKIDKELINRIRATQQRPHHFLRKGIFPSHRDMQVLDAYENK 83
 DB 2 SAKGIDYDKLIVRFSSSKIDKELINRIRATQQRPHHFLRKGIFPSHRDMQVLDAYENK 61
 QY 84 KPEFYLYTGRGPSSAMHNGHLIPFIETKWLQDVFNVPVLIQWTDDEKYLMDLTDQAVG 143
 DB 62 KPEFYLYTGRGPSSAMHNGHLIPFIETKWLQDVFNVPVLIQWTDDEKYLMDLTDQAVG 121
 QY 144 DAVENAKDIIACGPDINKTIFPSDDLYMGSSGFYKNVVKIQKAVTFNQVKIGFTDSD 203
 DB 122 DAVENAKDIIACGPDINKTIFPSDDLYMGSSGFYKNVVKIQKAVTFNQVKIGFTDSD 181
 QY 204 CIGKISFPAIQAPBSFNSFPQIFRDRDIOCLICALIDQDPYFMTDVAIRIGYKPA 263
 DB 182 CIGKISFPAIQAPBSFNSFPQIFRDRDIOCLICALIDQDPYFMTDVAIRIGYKPA 241
 QY 264 ILHSTFFPALQAGQKMSASDPNSSIFLTPDKOKIKTVNKAHSGGRDTEIEHROFGN 323
 DB 242 ILHSTFFPALQAGQKMSASDPNSSIFLTPDKOKIKTVNKAHSGGRDTEIEHROFGN 301
 QY 324 CDVDSFMVLTFFLEDDDKLEQIRKDYTGAMLTGLKKALEIVLOPLIAEHQARRKEVT 383
 DB 302 CDVDSFMVLTFFLEDDDKLEQIRKDYTGAMLTGLKKALEIVLOPLIAEHQARRKEVT 361
 QY 384 DEIVKEFMTPRKLSPDFQ 401
 DB 362 DEIVKEFMTPRKLSPDFQ 379

RESULT 10

AAE13494
 ID AAE13494 standard; Protein; 392 AA.

AC AAE13494;
 DT 12-FEB-2002 (first entry)

DE Human inactive tryptophanyl t-RNA synthetase in pET20B.

KW Human; tryptophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;
 KW TrpRS; vascular endothelial cell function; angiogenesis; wound healing;
 KW re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;
 KW diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;
 KW angiography; gene therapy; tumour; inflammation; vascular permeability;
 KW rheumatoid arthritis; psoriasis; diabetic retinopathy.

OS Homo sapiens.

PN WO200175078-A1.

PD 11-OCT-2001.

PF 21-MAR-2001; 2001WO-US08975.

PR 31-MAR-2000; 2000US-193471P.

PA (SCRI) SCRIPPS RES INST.

PI Schimmel P, Makasugi K;

DR WPI; 2002-010784/01.

DR N-PSDB; AAD22485.

PT Novel truncated tryptophanyl-tRNA synthetase polypeptides capable of
 PT regulating vascular endothelial cell function, preferably angiogenesis,
 PT is useful for treating solid tumor or suppressing tumor metastasis in
 PT mammal

PS Disclosure; Page 135-136; 149pp; English.

CC The patent discloses human aminoacyl tRNA synthetases, particularly

truncated tryptophanyl-tRNA synthetases (TrpRS) comprising a Rossmann fold nucleotide binding domain and polynucleotides encoding them. The invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS sequences are useful for regulating vascular endothelial cell function, preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound healing agents for re-vascularizing damaged tissues. They are useful for treating full-thickness wounds (e.g. dermal ulcers, including pressure sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS sequences can also be used in plastic surgery when reconstruction is required following a burn, other trauma, or even for cosmetic purposes. Angiogenic TrpRS is also used in association with surgery and following the repair of cuts, for promoting endothelialisation in vascular graft surgery and for repairing the damage of myocardial infarction and in conjunction with coronary bypass surgery by stimulating the growth of transplanted tissue. TrpRS is also used in conjunction with angiography. TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in immunoassays to detect the presence of tumours. They are also useful for blocking endogenous angiogenic activity and retard the growth of solid tumours. These antibodies may also be used to treat inflammation caused by increased vascular permeability. Inhibiting the activity of TrpRS by antisense technology is useful for preventing further growth or even regressing solid tumours, and for treating rheumatoid arthritis, psoriasis, diabetic retinopathy, all of which are characterised by abnormal angiogenesis. The present sequence is human inactive tryptophanyl t-RNA synthetase (TrpRS) in pET20B.

Sequence 392 AA:

Query Match 94.0%; Score 1988; DB 23; Length 392; Best Local Similarity 100.0%; Pred. No. 2.5e-196; Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

24 SAKGIDYDKLIVRGSSKIKELINLRATGQPHHFLARGFFSHRDMNVLDAAENK 83
 2 SAKGIDYDKLIVRGSSKIKELINLRATGQPHHFLARGFFSHRDMNVLDAAENK 61
 84 KPFYLYTGRGSSSAMVGHILPFIETKWLQDVNVNVLVQMTDEKYLKQDLTLDOAYG 143
 62 KPFYLYTGRGSSSAMVGHILPFIETKWLQDVNVNVLVQMTDEKYLKQDLTLDOAYG 121
 144 DAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKVVVKIOKHTVNOYKGIFFPTSD 203
 122 DAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKVVVKIOKHTVNOYKGIFFPTSD 181
 204 CIGISFPALQAAFSFNSFPQIFRDRDIOCLPCALDDPYFRMRDVAAPRIQYKPA 263
 182 CIGISFPALQAAFSFNSFPQIFRDRDIOCLPCALDDPYFRMRDVAAPRIQYKPA 241
 264 LHSSTFPALQAAQTAKASDPNSIFLTDPAKQIKRVNKHAFSGGRDTIEHRQFGCN 323
 242 LHSSTFPALQAAQTAKASDPNSIFLTDPAKQIKRVNKHAFSGGRDTIEHRQFGCN 301
 324 CDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARRKEVT 383
 302 CDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARRKEVT 361
 384 DEIVKEFMTPRKLSFDFQ 401
 362 DEIVKEFMTPRKLSFDFQ 379

RESULT 11
 ABB64621 standard; Protein; 430 AA.
 ID ABB64621;
 AC ABB64621;
 XX 26-MAR-2002 (first entry)
 DT 26-MAR-2002 (first entry)
 XX Drosophila melanogaster polypeptide SEQ ID NO 20655.
 DE Drosophila melanogaster polypeptide SEQ ID NO 20655.
 XX Drosophila melanogaster polypeptide SEQ ID NO 20655.
 KM pharmaceutical.
 KW pharmaceutical.

XX Drosophila melanogaster.
 OS Drosophila melanogaster.
 XX WO200171042-A2.
 PN 27-SEP-2001.
 XX 23-MAR-2001; 2001MO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PMD, Myers EW;
 PI WPI; 2001-656860/75.
 DR N-PSDB; ABL08724.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
 PT Disclousure; SEQ ID NO 20655; 21pp + Sequence Listing; English.
 PS The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABBS70272).
 CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 CC
 XX Sequence 430 AA:

Query Match 64.8%; Score 1370.5; DB 22; Length 430; Best Local Similarity 64.5%; Pred. No. 1.7e-132; Matches 253; Conservative 62; Mismatches 76; Indels 1; Gaps 1;

7 ATEAEDFVDPWTVVQSSAKGIDYDKLIVRGSSKIKELINLRATGQPHHFLARGI 66
 38 ATEAEDFVDPWTVVQSSAKGIDYDKLIVRGSSKIKELINLRATGQPHHFLARGI 97
 67 FFSHRDMNVLDAAENKPPYLYTGRGSSSAMVGHILPFIETKWLQDVNVNVLVQMT 126
 98 FFSHRDLHTLTLREQKPFYLYTGRGSSSAMVGHILPFIETKWLQDVNVNVLVQMT 157
 127 DDEKYLKQDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKVVVKIOK 186
 158 DDEKYLKQDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKVVVKIOK 217
 187 HTFENOVKGIFFGPTSDCIGISFPALQAAFSFNSFPQIFRDRDIOCLPCALDDPY 246
 218 HTFENOVKGIFFGPTSDCIGISFPALQAAFSFNSFPQIFRDRDIOCLPCALDDPY 276
 247 FRMTDVAAPRIQYKPAALHSSTFPALQAAQTAKASDPNSIFLTDPAKQIKRVNKHAF 306
 277 FRMTDVAAPRIQYKPAALHSSTFPALQAAQTAKASDPNSIFLTDPAKQIKRVNKHAF 336
 307 FSGGRDTIEHRQFGCNCDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKKALIE 366
 337 FSGGRDTIEHRQFGCNCDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKKALIE 396
 367 VLQPLIAEHQARRKEVTDEIVKEFMTPRKLSF 398
 397 TLPIVIEHQARRKEVTDEIVKEFMTPRKLSF 428

RESULT 12

[illegible]

PR	10-AUG-1999	99US-01463117
PR	11-AUG-1999	99US-01463119
PR	12-AUG-1999	99US-01463145
PR	13-AUG-1999	99US-01463541
PR	13-AUG-1999	99US-01463694
PR	16-AUG-1999	99US-01463368
PR	17-AUG-1999	99US-01464175
PR	18-AUG-1999	99US-01464926
PR	20-AUG-1999	99US-01463722
PR	20-AUG-1999	99US-01463723
PR	20-AUG-1999	99US-01463929
PR	23-AUG-1999	99US-01463902
PR	23-AUG-1999	99US-01463930
PR	23-AUG-1999	99US-01505660
PR	26-AUG-1999	99US-01500884
PR	27-AUG-1999	99US-01501085
PR	27-AUG-1999	99US-01510666
PR	27-AUG-1999	99US-01510800
PR	30-AUG-1999	99US-01513003
PR	31-AUG-1999	99US-01514338
PR	01-SEP-1999	99US-01519330
PR	07-SEP-1999	99US-01523363
PR	10-SEP-1999	99US-01530700
PR	13-SEP-1999	99US-01532758
PR	15-SEP-1999	99US-01540418
PR	16-SEP-1999	99US-01540403
PR	20-SEP-1999	99US-01544729
PR	22-SEP-1999	99US-01551339
PR	23-SEP-1999	99US-01555486
PR	24-SEP-1999	99US-01556569
PR	28-SEP-1999	99US-01564458
PR	29-SEP-1999	99US-01565696
PR	04-OCT-1999	99US-01571171
PR	05-OCT-1999	99US-01577653
PR	06-OCT-1999	99US-01578655
PR	07-OCT-1999	99US-01580229
PR	08-OCT-1999	99US-01582032
PR	12-OCT-1999	99US-01583369
PR	13-OCT-1999	99US-01582993
PR	13-OCT-1999	99US-01592984
PR	13-OCT-1999	99US-01593229
PR	14-OCT-1999	99US-01593329
PR	14-OCT-1999	99US-01593331
PR	14-OCT-1999	99US-01596377
PR	14-OCT-1999	99US-01596384
PR	18-OCT-1999	99US-01595984
PR	21-OCT-1999	99US-01607441
PR	21-OCT-1999	99US-01607677
PR	21-OCT-1999	99US-01607668
PR	21-OCT-1999	99US-01607700
PR	21-OCT-1999	99US-01608174
PR	21-OCT-1999	99US-01608815
PR	22-OCT-1999	99US-01609815
PR	22-OCT-1999	99US-01609881
PR	22-OCT-1999	99US-01610394
PR	25-OCT-1999	99US-01614005
PR	25-OCT-1999	99US-01614006
PR	25-OCT-1999	99US-01613559
PR	26-OCT-1999	99US-01613360
PR	26-OCT-1999	99US-01613361
PR	28-OCT-1999	99US-01613621
PR	28-OCT-1999	99US-01613922
PR	28-OCT-1999	99US-01614923
PR	29-OCT-1999	99US-01621442

```

Query Match      61.6%; Score 1304; DB 21; Length 402;
Beet Local Similarity 60.3%; Pred. No. 1.le-125;
Matches 240; Conservative 67; Mismatches 87; Indels 4; Gaps 2.

Oy      6 DATEAE--EDFVDPTVQTSSAKCIDYDKLIVRPGSSKIDKELINRIERATGQRPHEFLR 63
      | ||| | | ||| | |||| | || :||: ||: ||: ||: || |||

```


PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147483.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148664.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160960.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.

Query Match 61.6%; Score 1304; DB 21; Length 426;
Best Local Similarity 60.3%; Pred. No. 1.2e-125;
Matches 240; Conservative 67; Mismatches 87; Indels 4; Gaps 2;

QY 6 DATEAR-EDFVDPMTVQVSSAKGIDYDKLIRFGSSKIDKELINIRERATGQRPHHFR 63
DB 31 DEREASSQVNVNPMVWSAKDGCKIDYDKLIRFGQRLDESILIDVOKLITSRQPHVFLR 90

QY 64 RGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVNPLVI 123
DB 91 RSVFFAHRDFNELIDAYERGDKPFYLYTGRGSSSEALHGLHLPMTKYLQEAFKVPLVI 150

QY 124 QMTDEKYLKDOIITLDQAYGDAVENAKDIIACGPDINKTFTISDDLDMGMSGGFYKNVYK 183
DB 151 QLTDEKSIWKNLISVEESQRLARENKDOIACGPDVTKTFTISDPDYVG--GAFYKNVYK 208

QY 184 IQKHTFNGVKGIFGFTSDSCIGKISFPALQAPSPNSFPQIFRDRTOICLIPCAIDQ 243
DB 209 VGKCVTLNKGMPGSGEDPIAKLSFPVQAVPSFSPSPHLPCKDKMLRCLIPCAIDQ 268

QY 244 DPFYFMTRDVAPRIGYKPKPALHSTFFPALQAGQTMASDPNSIFLTDTAQOIKTKVN 303
DB 269 DPFYFMTRDVAPRIGYKPKPALHSTFFPALQAGQTMASDPNSIFLTDTAQOIKTKVN 328

QY 304 KNAFSGGRPTIEBROFGNCVDVNSFWMLTFFLEDDDLQRIKDYTGAMLTGLKVA 363
DB 329 RYAFSGGQDSIEKHRELGNLEVIDIPVKYLSFPLEDDSELEHIXKEYGGRMLTGEVKR 388

QY 364 LIEVLQPLAEHQARKEVTDIEVKEFMTPKLSFDFQ 401
DB 389 LIEVLTEIVEKRRAPAAVTDENVDAFMAVRPLPSKFE 426

RESULT 15
AAB66931
ID AAB66931 standard; Protein; 424 AA.
XX AAB66931;
AC AAB66931;
XX 12-APR-2001 (first entry)
DT 12-APR-2001 (first entry)
XX Tryptophanyl-tRNA synthetase.
DE Tryptophanyl-tRNA synthetase.
XX Tryptophanyl-tRNA synthetase; enzyme.
KW Tryptophanyl-tRNA synthetase; enzyme.
OS Candida albicans.
XX US6174713-B1.
PN US6174713-B1.
XX 16-JAN-2001.
PD 16-JAN-2001.
XX 16-JUN-1997; 97US-0876685.
PF 16-JUN-1997; 97US-0876685.
XX 16-JUN-1997; 97US-0876685.
PR (CUBI-) CUBIST PHARM INC.
PA Shen X, Houman F;
XX WPI, 2001-201806/20.
DR N-PSDB; AAF55855.
XX

PT New nucleic acids encoding Candida cytoplasmic tryptophanyl-tRNA
 PT (ctRNA) synthetases, useful for producing recombinant ctRNA synthetases
 PT and detecting inhibitor of Candida ctRNA synthetase function -
 XX
 XX

PS Claim 4; Fig 1; 32pp; English.

CC The present sequence is Candida albicans cytoplasmic tryptophanyl-tRNA
 CC (ctRNA) synthetase. ctRNA synthetase is useful for producing recombinant
 CC ctRNA synthetases and detecting inhibitors of ctRNA synthetase.
 XX

Sequence 424 AA;

Query Match 57.6%; Score 1218.5; DB 22; Length 424;
 Best Local Similarity 57.1%; Pred. No. 8.4e-117;
 Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

```

QY      8 TEAEDFDVDPWTVQ---TSSAKIDYDYLIVFGSSKIDKELINRIERATGQRPHELR 63
DB      13 TESESEOKITPWEVEGAVVDGKSMGIDYDKLISQFTKHITETLERFKQVTGEPPHFLK 72
QY      64 RGIFFSHRDMNGVLDAYENKKPPLYLTGRGPSEAMHVGHLIPITFKMLQDVENVPLVI 123
DB      73 RGVFFSQKDLRLIDLYEHGEPPFLYTGKSPSSDMLGHMVPFLFTKMLQEVFVPLVI 132
QY      124 QMTDDEKYLWK-DLTLDQAYGDAVENAKDIIACGFDINKFTIFSDDVYMGSSGFYKNV 182
DB      133 ELTDEKFLFKQLITDDVKGFAENAKDIIAVGFNPENTFIFSDLQVMG--GAFYENVV 190
QY      183 KIQKAVTFNOVKGIFGFTDSDCIKISPAIQAPSFNSFPQIFRRTDIQCLIPCAID 242
DB      191 RTSROITSTAKAVFGFTDSDCIKIHFAISIQIATAFPSSFPDVLGLPPTKPLIPCAID 250
QY      243 QDPYFRMTRDVAPRIGYPRKALHSTFFPALOGAOTKMSASDPNSSIFLDTAKQIKTKV 302
DB      251 QDPYFRVCDVDADKLFTRPALIHAKFPALOGASTKMSASDPTTTSIFMGDTAKQIQKKI 310
QY      303 NGHAFSGGRDTIEHRQFGCNDVDSFWYLTFLEDDDKLEQIRKDYTSGAMLTGELTK 362
DB      311 NKYAFSGGRATAEHRHELGNPEVDVAFQYLSFFSYDDEKLAQLEGGYRKGEILSGEMCK 370
QY      363 ALIEVLOPLIAHQARKKEVTDIYKEFMTPRKLSF 398
DB      371 ECITVLOEFVSAYQERRSKVDQVVEKFMKPHKLVF 406

```

Search completed: July 10, 2003, 12:30:12
 Job time : 33.08 secs

